PATENTTI- JA REKISTERIHALLI
NATIONAL BOARD OF PATENTS A EGISTRATION

T / F I 0 0 / 0 0 2 4 9

Helsinki 22.3.2000

09/936823

REC'D 19 APR 2000
WIPO PCT

E T U O I K E U S T O D I S T U S
P R I O R I T Y D O C U M E N T

Hakija Applicant	**Valtion teknillinen tutkimuskeskus** **Espoo**
Patenttihakemus nro Patent application no	**990667**
Tekemispäivä Filing date	**25.03.1999**
Kansainvälinen luokka International class	**C07K**

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Keksinnön nimitys
Title of invention

"Process for partitioning of molecules"
(Menetelmä molekyylien erottamiseksi)

Täten todistetaan, että oheiset asiakirjat ovat tarkkoja jäljennöksiä patentti- ja rekisterihallitukselle alkuaan annetuista selityksestä, patenttivaatimuksista ja piirustuksista.

This is to certify that the annexed documents are true copies of the description, claims and drawings originally filed with the Finnish Patent Office.

Pirjo Kalla
Tutkimussihteeri

Maksu 300,- mk
Fee 300,- FIM

Osoite: Arkadiankatu 6 A
P.O.Box 1160
FIN-00101 Helsinki, FINLAND
Puhelin: 09 6939 500
Telephone: + 358 9 6939 500
Telefax: 09 6939 5328
Telefax: + 358 9 6939 5328

Process for partitioning of molecules

Field of the invention

The present invention relates to isolation and purification of proteins in aqueous two–phase systems (ATPS). Specifically the invention provides processes for partitioning of molecules of interest in ATPS by fusing said molecules to targeting proteins which have the ability of carrying said molecule into one of the phases.

Background of the invention

Liquid–liquid extraction in an aqueous two–phase system (ATPS) can offer a powerful technique for isolation and purification of proteins. The separation of macromolecules and particles by means of liquid–liquid extraction is well known (Albertsson, 1986; Walter *et al.*, 1985; Kula, 1990). Mainly polyethylene glycol (PEG) – salt, PEG–dextran and PEG–starch systems have been in use. More recently detergents and detergents with reversed solubility were discovered as suitable methods for separation of macromolecules, and especially for the separation of proteins.

An advantage of aqueous two–phase systems (ATPS) is that they are especially suited for large scale processing of microbial proteins not only from culture supernatants but also from crude extracts containing cells and cell debris (Kula, 1979; Kula, 1985). Characteristic features of biological fluids as well as suspensions are rather small particle sizes, low density differences between fluid and suspended solids, high viscosities of the extracts and high compressibility of the solids (Hustedt *et al.*, 1985; Bender and Koglin, 1986). These attributes decrease the performance of conventional methods for solid–liquid separation like centrifugation and filtration at the beginning of a protein recovery process. Using an aqueous two–phase system removal of solids can be integrated into a liquid–liquid separation step, clarification is thus combined with an initial purification (Kula, 1979; Kula, 1985).

After the extraction process phase separation can be accomplished by settling under gravity as well as by centrifugation (Kula, 1985). ATPS can be applied in various scales from very small laboratory scale to large industrial scale thus suiting for various proteins, purposes and needs. With regard to industrial purposes commercially available centrifugal separators can be used to shorten separation time. Several authors have investigated the potential of centrifugal separators of various design for processing of large volumes of aqueous two phase systems (Kula, 1979; Kula *et al.*, 1981, Kula *et al.*, 1982; Kula, 1985). In these studies the authors have used polymer/polymer or polymer/salt systems and the results of these investigations demonstrate the feasibility of continuous separation of aqueous two-phase systems in centrifugal separators.

Extraction systems based on nonionic surfactants have been described as an alternative to standard polymer/polymer or polymer/salt systems. Phase forming surfactants are e.g. polyoxyethylene type nonionic detergents. The basis of this type of aqueous two-phase system is the temperature-dependent reversible hydration of the polar ethylene oxide head groups. The temperature at which the phase separation occurs is referred to as the cloud-point (cloud-point extraction). This kind of aqueous two-phase system is especially suited for the extraction of amphiphilic biomolecules. The potential of this type of two-phase system for separating membrane bound proteins from cytosolic and peripheral membrane proteins was first demonstrated by Bordier (1981). Heusch and Kopp (1988) have been able to demonstrate that lamellar structures formed in the miscibility gaps of polyglycol ether / water systems are responsible for the selective extraction of hydrophobic substances.

Recently, the successful application of a surfactant-based aqueous two-phase system for the extraction of a membrane bound protein (cholesterol oxidase) from the unclarified culture medium of the gram-positive microorganism *Nocardia rhodochrous* on a bench scale has been reported (Minuth *et al.*, 1995). By addition of only one chemical compound a product release through solubilization was possible in homogeneous phase and in a second step a clarification as well as an initial purification was achieved by an extraction process at elevated temperatures separating the detergent rich phase. A closed concept was further developed for the production of the membrane bound enzyme by surfactant-based extraction, organic solvent extraction and anion-

exchange chromatography, which gave a product suitable for analytical applications (Minuth *et al.*, 1996).

In aqueous two phase systems the desired target e.g. a protein should partition selectively into one phase (preferentially the lighter phase) while the other substances should partition into the other phase (preferentially the heavier phase). In PEG/salt and PEG/dextran and similar systems there are several driving forces for a substance like charges, hydrophobic, hydrophilic forces or the dependence on conformation or ligand interaction (Albertsson, 1986). The forces leading to separation in detergent based aqueous two phase systems are suggested to be primarily hydrophobic (Terstappen *et al.*, 1993). Even if a lot of work has been carried out in the field of prediction in ATPS, none of the designed models provides a physical picture of the phase behaviour and prediction is hardly possible (Johansson *et al.*, 1998).

In ATPS the partitioning coefficient is defined as the concentration (activity in case of an enzyme) of the target in the top phase divided by the concentration (enzyme:activity) of the target protein in the bottom phase.

$$K=\frac{c_{i,T}}{c_{i,B}}$$

Yield: is defined as the amount of target in the top phase divided by the sum of the amount of target in top and bottom. This leads to the following equation

$$Y_T=\frac{1}{1+\left[\frac{V_B}{V_T}\bullet\frac{1}{K}\right]}$$

If the desired substance is directed to the heavier phase (as it can be the case using Triton) the yield is defined by

$$Y_B=\frac{1}{1+\left[\frac{V_T}{V_B}\cdot K\right]}$$

The volume ratio of the two coexisting phases are defined by the volumes of the lighter over the heavier phase respectively.

$$R = \frac{V_T}{V_B}$$

An example of useful proteins facing problems in purification in a cost-effective way are the commonly used industrial enzymes used as biocatalysts, the glycosyl hydrolases, proteases and lipases produced by fungi and bacteria. These are used in e.g. laundry, textile, paper and pulp, food and feed industry. The fact that microbes produce many different enzymes during their growth and the fact that some of these may be undesired in certain applications leads to a need to enrich the active component(s). This enrichment can be performed by choosing appropriate growth conditions, by genetic engineering and/or by down-stream processing (e.g. purification of the active component(s)).

Purification of proteins are generally performed by chromatography. Usually gel-chromatographic methods are used based on ion-exchange, hydrophobic interaction, affinity chromatography and molecular sieving. Methods like electrophoresis and crystallisation can also be used. These methods are well known in the art and suitable for proteins of fairly high market value. In case of bulky enzyme production these methods, however, are too expensive in order to keep the final product on a compatible price level. Due to similar properties of these enzymes several purification steps are usually needed to separate the proteins from each other. This often causes low final yields and therefore a high loss of product.

Many extracellular hydrolases produced by the filamentous fungus *Trichoderma* are currently used in different industrial applications in large scale. These hydrolases are e.g. hemicellulases (such as xylanases and mannanases), cellulases (such as endoglucanases and cellobiohydrolases) and proteases. Purification of these is well known in the art (Bhikhabhai *et al.*, 1984; Pere *et al.*, 1995), but for large industrial applications the purification methods are too expensive. Alternative methods to enrich these hydrolases have been used, including deletion of undesired genes by genetic engineering (Suominen *et al.*, 1992). However, even after extensive genetic engineering some minor undesired activities may still be present in the final product.

ATPS have been studied in purification of cellulases of *T. reesei* and the purification of an endoglucanase III showed some promising results, enriching the yield of the protein in the upper phase (U.S. Pat. 5,139,943). ATPS have also been studied in purification of lipases, endoxylanase and natamycin (EP 0 574 050 A1). No K and Y values are, however, mentioned.

As in other protein purification methods, similar properties of proteins produced by an organism impair also in ATPS, e.g. selective separation of one protein is not achieved optimally. To obtain selectivity in purification affinity chomatographic methods are used especially for analytic purposes and in purification of high-value products. These include immunoaffinity chromatography and various fusion protein strategies well described in the art such as fusing the protein of interest to an other protein (e.g. glutathione-S-transferase), protein domain (e.g. protein A-ZZ domain) or small peptide (e.g. His-tag), which selectively bind to the solid carrier and thus the recovery of the fusion partner is obtained as well. The fusion protein can be suitable for the particular purpose as such or cleavage of the product from the added fusion partner may be desired. There are well-known methods in the art on cleavage of fusion proteins from their partners by proteases, e.g. by factor X, thrombin or papain or by genetically introducing a protease cleavage site (e.g. Kex2 site) or autoprosessing domains (e.g. Intein, New England Biolabs) or by chemical cleavage.

ATPS offers advantages mainly with respect to technology compared with the solid state based separation systems e.g. affinity column-based techniques. The scale-up of extractive enzymes is relatively simple utilising commercially available equipment and machinery common in the chemical industry. In addition, it can be used in a continuous process and it can be relatively cost extensive. It can be used as a single step for clarification, concentration and purification. ATPS can be used as a first capture-step, but for bulk products often no further purification is needed.

To aid selective separation in two-phase systems, recent publications have described the fusion of small peptide tags of 12 amino acids to the protein to be purified. The most successful of these soluble peptides are containing tryptophans. So far they have mainly

be applied for very small molecules like the staphylococcal protein A derivative ZZT0 (Berggren *et al.*, 1999; Hassinen *et al.*, 1994; Köhler *et al.*, 1991).

Use of ATPS has so far been limited to certain targets. Due to the advantages of ATPS in protein separation, purification and localisation, highly selective and powerful methods should be developed. This is especially important for large scale processes where ATPS in general is very inexpensive as a first capture step or as the only step for purification, clarification and concentration.The system should be universal so that the technique would be strong enough to mediate separation of in principle any component to the desired phase irrespective of its size or biochemical properties.

Description of the invention

In this invention we describe selective separation and partitioning of molecules and particles fusing them with targeting proteins having the capability to carry the molecule or particle of interest to the desired phase in ATPS, and to keep it in this phase if wanted. This invention is directed to make ATPS usable for every biotechnological product. By addition of the targeting protein to selective products, either by genetic tagging of proteins, by chemical binding, glueing or by use of any other technique, the product molecule can be turned more suitable for separation in ATPS. Using ATPS the product or certain component is therefore driven to one phase while the other components or by–products are directed to the other phase(s).

We also describe that efficient separation in ATPS can be obtained using targeting proteins which are/can be larger than the described small soluble synthetic peptide tags of 12 amino acids or less. These targeting molecules can aid in separating of small molecules but even large proteins and particles. Unlike the small peptide tags, it is not necessary that they contain tryptophan residues, although they may do so. They can be hydrophobic or moderately hydrophobic and/or amphipathic in nature, either in monomeric form or when forming aggregates. Such proteins can be found in nature or they can be designed, or obtained through for instance methods known in art for mutant generation, gene shuffling or directed evolution. Suitable targeting molecules can be

screened for instance by fusion the product of interest to a library of natural or mutant sequences, and screening the ability of the fusion molecules to separate in ATPS.

Examples of molecules suited as targeting proteins in ATPS found in nature are hydrophobin-like small proteins. Hydrophobins are secreted proteins with interesting physico-chemical properties that have recently been discoverd from filamentous fungi (Wessels, 1994; Wösten and Wessels, 1997; Kershaw and Talbot, 1998). One characteristic feature of these proteins is their moderate hydrophobicity. They are usually small proteins, approximately 70 to 160 amino acids, containing eight cysteine residues in conserved pattern, and do usually not contain tryptophanes. However, also multimodular proteins with one or several hydrophobin domains and e.g. proline-rich or asparagine/glycine repeats, or hydrophobins containing less than eight cysteine residues have been characterized (Lora *et al.*, 1994; Lora *et al.*, 1995; Arntz and Tudzynski, 1997). Hydrophobins have been divided into two classes based on their hydropathy profiles (Wessels, 1994).

Today most protein data exists for the hydrophobins Sc3p of *Schizophyllum commune* (class I), and cerato-ulmin of *Ophiostoma ulmi* and cryparin of *Cryponectria parasitica* (class II), although more than 30 gene sequences for hydrophobins have been published (Wösten and Wessels, 1997). HFB genes are often naturally highly expressed but due to special requirements in cultivation conditions and the biochemical properties of the proteins, purification of HFBs in large amounts have been difficult. For instance only relatively moderate production levels of a few mg per liter of Sc3 hydrophobin in static cultures are obtained (Han Wösten, personal communication). Published purification procedures include e.g. multi-step extraction from fungal cell walls using organic solvents and bubbling or freezing of cuture filtrates (Wessels, 1994). No reports of successful production of hydrophobins are available; levels of cerato-ulmin were no higher than those obtained with other naturally occuring fungal isolates (Temple *et al.*, 1997).

Upon shaking hydrophobin-containing solutions, the protein monomers form rodlet-like aggregates. These structures are similar to the ones found on surfaces of aerial stuctures. The self-assembly of purified Sc3 hydrophobin into a 10 nm thick amphipatic layer on

hydrophilic and hydrophobic surfaces has been demonstrated (Wösten *et al.*, 1993; Wösten *et al.*, 1994b). This film is very strongly attached to the surface and not broken, for instance, by hot detergent. The hydrophobic side of the layer on hydrophilic surfaces shows properties similar to teflon (Wessels, 1994). The Sc3 assemblages, as well as those of cerato-ulmin and cryparin, also form on gas-liquid or gas-air interphases thus stabilizing air bubbles or oil droplets in water.

Surface activity of proteins is generally low but hydrophobins belong to surface-active molecules their surfactant capacity being at least similar to traditional biosurfactants such as glycolipids, lipopetides/lipoproteins, phospholipids, neutral lipids and fatty acids (Wösten and Wessels, 1997). In fact Sc3 hydrophobin is the most potent biosurfactant known. It lowers the water surface tension to 24 mJm^2 at a concentration of 50 μg/ml due to a conformational change during self-assembly of monomers into an amphipathic film (Wösten and Wessels, 1997).

Hydrophobin-like molecules, vary in their properties. For instance, rodlet-forming capacity has not been assigned for all hydrophobins (such as some class II), or they might have a weaker tendency to form stable aggregates (Russo *et al.*, 1982; Carpenter *et al.*, 1992). Another group of fungal amphiphatic proteins are repellents (Wösten *et al.*, 1996 (Ustilago), for review, see Kershaw and Talbot, 1998). Other type of proteins suited as targeting proteins for ATPS may consequently have only some of the features assigned to hydrohobins. Other suitable proteins are hydrophobic ones such as e.g. lipases, cholesterol oxidase, membrane proteins, small peptide drugs, aggregating cell wall proteins, lipopetides or any parts of these or combinations of these, and other molecules like glycolipids, phospholipids, neutral lipids, fatty acids in combination with proteins or peptides.

In this invention the targeting protein, such as a hydrophobin-like protein or parts of it is bound to the product molecule or the component to be separated. First, phase forming materials and eventually possibly also additional salts are added to a watery solution containing the fusion molecule or component and optionally also the contaminating materials. The added agents are mixed to facilitate their solubilisation. As soon as they are solubilized the two phases are formed either by gravity settling or centrifugation. In

the separation the target protein drives the product to for instance the detergent-rich phase which could either be the top or the bottom phase. The method is not only useful for purification of products of intrest but also for keeping the product or the component of interest, such as a biocatalyst, in the particular phase which enables certain useful biotechnical reactions.

Several ATPS systems are suitable for performing this invention. These include PEG containing systems, detergent based systems and novel thermoseparating polymers. Detergent based systems can be nonionic, zwitterionic, anionic or kationic. The system can be based on amphiphilic polymeric detergents, micelle forming polymers. Novel polymers can be based on polyethylene-polypropylene copolymers such as pluronic block copolymers, Brij, polyoxyethylene derivatives of partial ethers of fatty acids made by adding polyoxyethylene chains to the nonesterified body and polyoxyethylene derivatives. The well known PEG/salt, PEG/dextran and PEG/starch (or derivatives such as Reppal, hydroxipropyl starch) systems where PEG and water are forming the top phase and dextran/starch/salt and water are forming the bottom phase. As salts are used phosphate, citrate, sulfate or others. In the present process the target is partitioning mainly to the top phase, while most of the contaminants are separating mainly to the bottom phase. Some hydrophobic contaminants might partition to the top phase as well. Using detergent based systems only one phase forming detergent has to be added. Optionally, salts and other chemicals can be used in addition. The mentioned chemicals are added, and the solution is mixed. After mixing the separation can take place either by centrifugation or gravity settling. In order to separate into two phases the temperature of the solution has to be over the cloud point of the detergent. The solution has to be heated if the cloud-point is not reached otherwise. If wanted, a second separation step can follow after a first extraction step and the product rich phase can be further purified. Also the remaining product in the product poor and by-product rich phase can be re-extracted. Very good K values can be obtained and the yields and concentration factors are high.

The process of the present invention can advantageously be used in laboratory scale but is especially suitable for large scale separations. It can successfully be used in the separation of proteins and components from large fermentations. Using genetic

modifications, the method can be used to purify any protein of interest including extracellular enzymes and proteins such as cellulases and hemicellulases from mixtures containing large amounts of protein such as several garms per liter. Furthermore, this separation can be obtained from various culture media including industrial media containing particular materials such as cellulose and spent grain. The separation can be done directly from the fermentation broth which can additionally contain cells, even viscous filamentous fungi. High biomass levels are acceptable for the process as explained in example 9. An example is the extracellular endoglucanase I from the fungus *Trichoderma reesei* which can be tagged for instance with the class 2 HFBI and can for example be separated with the nonionic polyoxyethylene C12–C18EO5. In this example the detergent rich phase is the lighter phase and contains most of the tagged endoglucanase while most of the other cellulases, proteases and other enzymes remain in the heavier phase. The mycelium separates to the bottom phase, too. The separation can be achieved using separation temperatures higher than 25°C. The temperature can be decreased if certain salts like NaCl or K_2SO_4 are added.

The invention describes separation of molecules produced in various different organisms such as bacteria, yeast and filamentous fungi. The invention is suitable for purification of product molecules from extra– or intracellular locations, including cell wall bound molecules. It provides examples how the fusion molecule can be secreted by these different organisms but also provide an example how the fusion can be produced intracellularly.

The invention further describes how fusion molecules consisting of several domains can be constructed and successfully expressed and produced. The invention describes fusions of the targeting molecule to a small protein (CBD), to a moderately sized protein (EGI) and to a huge highly glycosylated protein (FloI), and different domain variations of these. These molecules can be ready as such for biotechnical use. Alternatively the product can be cleaved from the targeting protein by any method known in the art such as with proteases e.g. thrombin, factor X, papain or by chemical cleavage. Furthermore, ATPS is a preferential means to be used to separate the product from the targeting protein after cleavage, or these can be separated with other methods known in the art.

A suprising feature is that the targeting protein can also be used to carry large particles to the desired phase in ATPS. This can be obtained if the particles already contain proteins suited for targeting such as spores/conidia do in case of fungi. The targeting protein can also be attached to the particles or compounds *in vitro*. If cells are separated, the targeting protein can alternatively be expressed in the recombinant cells in such a way that it is exposed at the cell surface whereby it mediates the separtion of the cells in ATPS. A teaching how this can be done is provided in example 22. Other types of molecules which direct the targeting molecule to the cell surface can be found e.g. in the literature including bacterial outer membrane proteins and lipoproteins (Ståhl, S. and Uhlén, M. (1997) Trends Biotechnol. 15: 185–192), and yeast proteins α–agglutinin and flocculin (Schreuder, M.P. et al (1996) Trends Biotechnol. 14: 115–120; Klis , F.M. et al. (1994) WO 94/01567; Frenken, L.G. (1994) WO 94/18330).

A further advantage of the system is that the invention combined with ATPS provides a means to separate the product or desired component not only from other unnecessary or unwanted proteins but also from harmful proteins such as proteases as described in example 6. Thus, the invention is particularly suited for production and purification of heterologous proteins, e.g. sensitive mammalian proteins usually produced in limited amounts in heterologous hosts. Such proteins are for instance antibodies or fragments thereof, interferon, interleukin, oxidative enzymes and any foreign protein which can otherwise be produced in the host. It is possible that separation of the product from e.g. culture medium can also be obtained on–line or semi–continously thus minimising the effect of proteases or other harmful components present in the culture. When produced intracellularly, the invention also provides means to separate the heterologous product for instance the inclusion bodies it may form from the cellular extracts.

This invention describes for the first time that fusion proteins containing hydrophobin–like molecules can be made and produced in significant amounts despite the very particular properties of hydrohobin–like molecules. Importantly, this invention also describes how recombinant strains producing increased amounts of hydrophobin–like proteins as such can be made. This provides a mean to produce the targeting protein for

uses in which it is wanted that the targeting protein is bound to the product or particle *in vitro*, to enable further separation of such molecules or particles in ATPS.

Importantly, this invention also decsribes how hydrophobin-like molecules can be purified in ATPS very efficiently. The molecules can be separated in the same way as the above mentioned fusions, for instance by PEG systems of by detergent based systems. Separation can be done from the culture medium or from cells. This provides a significant improvement in making pure preparates containing hydrophobin-like molecules since due to their properties their purification is very complicated and results in losses with the previously reported techniques as described above.

The invention is further illustrated by the following Examples which describe construction of the fusion molecules of the invention, and partitioning of the molecules of interest using the process according to the invention.

EXAMPLES

Example 1

Construction of vectors for expression of EGI and EGIcore HFBI fusion proteins under the *cbh1* and *gpd1* promoters of *Trichoderma* and *gpdA* promoter of *Aspergillus*

For construction of an EGI-HFBI fusion protein, *hfb1* (SEQ ID 1) coding region (from Ser-23 to the STOP codon) and a peptide linker (Val Pro Arg Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly) preceeding it was amplified with PCR using pTNS9 as a template and as a 5' primer **TCG GG C ACT ACG TG C CAG TAT AGC AAC GAC TAC TAC TCG CAA TGC *CTT GTT CCG CGT GGC TCT*** AGT TCT GGA ACC GCA (SEQ ID 2) and as a 3' primer TCG TAC GGA TCC TCA AGC ACC GAC GGC GGT (SEQ ID 3). pTNS9 has been decribed in detail in Example 19. The sequence in bold in the 5' primer encodes 16 C-terminal residues of EGI. The sequence in italics is a thrombin cleavage site and the underlined CACTACGTG is a DraIII site. The underlined GGATCC in the 3' primer is a BamHI site. The 280 bp PCR fragment was purified from agarose gel and ligated to pGEM-T T/A vector (Promega) resulting in pMQ102.

For construction of an EGIcore-HFBI fusion protein, the *hfb1* coding region (as above) was amplified with PCR using pTNS9 as a template and as a 5' primer **ACT ACA CGG AG G AGC TC G ACG ACT TCG AGC AGC CCG AGC TGC ACG CAG** AGC AAC GGC AAC GGC (SEQ ID 4) and as a 3' primer SEQ ID 3. The sequence in bold in the 5' primer encodes amino acids 410-425 in EGI and the underlined GAGCTC is a SacI site. The 260 bp PCR fragment was purified from agarose gel and ligated to pPCRII T/A vector (Invitrogen) resulting in pMQ111.

In the next step *Trichoderma* expression vectors for production of EGI-HFBI and EGIcore-HFBI fusion proteins under the control of *cbh1* promoter and terminator sequences were constructed. The expression vector used as a backbone in the constructs is pPLE3 (Nakari et al. (1994) WO 94/04673) which contains a pUC18 backbone, and carries the *cbh1* promoter (SEQ ID 5) inserted at the EcoRI site. The *cbh1* promoter is operably linked to the full length *egl1* cDNA (SEQ ID 6) coding sequence and to the *cbh1* transcriptional terminator (SEQ ID 7). The plasmid pMQ102 was digested with DraIII and BamHI and the released 280 bp fragment containing *hfb1* and linker sequences was purified from agarose gel and ligated to pPLE3 digested with DraIII and BamHI. The plasmid pMQ111 was digested with SacI and BamHI and the 260 bp fragment containing the *hfb1*

sequence was ligated to pPLE3 digested with SacI and BamHI. The resulting plasmids pMQ103 (Figure 1) and pMQ113 (Figure 2) carry the coding sequences for full-length EGI linked to HFBI via a peptide linker and for EGIcore linked to HFBI via its own linker region, respectively, under the control of *cbh1* promoter and terminator sequences.

Trichoderma expression vectors for production of EGI-HFBI and EGIcore-HFBI fusion proteins under the control of *gpd1* promoter and terminator sequences of *Trichoderma* and *gpdA* promoter and *trpC* terminator sequences of *Aspergillus* were constructed as follows. A SacII site was inserted inbetween the XbaI and PacI sites of pMV4 using as an adapter annealed primers TAA CCG CGG T (SEQ ID 8) and CTA GAC CGC GGT TAA T (SEQ ID 9). The resulting plasmid is pMVQ. pMV4 contains a pNEB193 (New England Biolabs) backbone, and carries a 1.2 kb *Trichoderma gpd1* promoter (SEQ ID 10) and a 1.1 kb *gpd1* terminator (SEQ ID 11) inserted at SalI-XbaI and BamHI-AscI sites, respectively. The expression casettes for EGI-HFBI and EGIcore-HFBI were released from pMQ103 and pMQ113 with SacII and BamHI, purified from agarose gel and ligated to pMVQ cut with SacII and BamHI. The resulting plasmids pMQ104 (Figure 3) and pMQ114 (Figure 4) carry the EGI-HFBI and EGIcore-HFBI casettes, respectively, under the control of *Trichoderma gpd1* transcriptional control sequences. Expression plasmids pMQ105 (Figure 5) and pMQ115 (Figure 6) containing EGI-HFBI and EGIcore-HFBI casettes, respectively, operably linked to the *gpdA* promoter and *trpC* terminator of *Aspergillus* were constructed. EGI-HFBI and EGIcore-HFBI casettes were released from plasmids pMQ104 and pMQ114 with XbaI and BamHI, blunted with T4 DNA polymerase and ligated to NcoI digesteded and T4 DNA polymerase treated pAN52-1 (SEQ ID 12). pAN52-1 contains a pUC18 backbone, and carries a 2.3 kb *gpdA* promoter and a 0.7 kb *trpC* terminator sequences of *A. nidulans*.

Example 2

Construction of vectors for over-production of HFBI on cellulase-inducing and -repressing media

For over-expression of HFBI under *cbh1* promoter the protein coding region of *hfb1* was amplified with PCR using as a template pEA10 (Nakari-Setälä et al. Eur. J. Biochem. (1996) 235:248-255). pEA10 carries a 5.8 kb genomic SalI fragment containing *hfb1* coding and flanking sequences. **GTC AA C CGC GG A CTG CGC ATC** ATG AAG TTC TTC GCC ATC (SEQ ID 13) was used as a 5' primer in the PCR and as a 3' primer SEQ ID 3. The sequence in bold in the 5' primer is 21 bp of

cbh1 promoter adjacent to translational start site of the corresponding gene and the underlined CCGCGG is a KspI site. The obtained fragment of 430 bp was digested with KspI and BamHI and ligated to pMQ103 digested with KspI and BamHI. The resulting plasmid pMQ121 (Figure 7) carries the coding sequence of *hfb1* operably linked to *cbh1* transcriptional control sequences. pEA10 plasmid is used for over-production of HFBI in cellulase-repressing conditions.

Example 3

Transformation of *Trichoderma* and purification of the EGI-HFBI and EGIcore-HFBI producing and HFBI over-producing clones

Trichoderma reesei strains QM9414 (VTT-D-74075) and Rut-C30 (VTT-D-86271) were co-transformed essentially as described (Penttilä et al., Gene (1987) 61:155-164) using 3-13 µg of the plasmids pMQ103, pMQ113, pMQ104, pMQ114, pMQ105, pMQ115, pMQ121 and pEA10 and as the selection plasmids 1-3 µg pToC202, p3SR2 or pARO21. pToC202 (pUC19 backbone) and p3SR2 (pBR322 backbone) plasmids carry 2.7 kb XbaI and 5 kb EcoRI-SalI genomic fragments of *A. nidulans*, respectively, containing the *amdS* gene (Hynes et al. Mol. Cell Biol. (1983) 3:1430-1439; Tilburn et al. Gene (1983) 26:205-221). pARO21 is essentially the same as pRLMex30 (Mach et al. 1994) and carries the *E. coli hph* gene operably linked to 730 bp of *pki1* promoter and 1 kb of *cbh2* terminator sequences of *T. reesei*. The Amd+ and Hyg+ transformants obtained were streaked three times onto plates containing acetamide and hygromycin, respectively (Penttilä et al. (1987) Gene 61:155-164). Thereafter spore suspensions were made from transformants grown on Potato Dextrose agar (Difco).

The production of the fusion proteins EGI-HFBI and EGIcore-HFBI and HFBI was tested by slot blotting or Western analysis with EGI and HFBI specific antibodies from shake flask or microtiter plate cultivations carried out in mimal medium supplemented with either glucose, lactose or a mixture of solka floc cellulose and/or spent grain and/or whey. The spore suspensions of the fusion protein producing clones were purified to single spore cultures on selection plates (containing either acetamide or hygromycin). To determine the best producers, production of the fusion proteins was analyzed again from these purified clones as described above.

T. reesei strains selected for further fermentor cultivations are VTT-D-98692 (pEA10), VTT-D-98492 (pMQ121), VTT-D-98693 (pMQ103), VTT-D-98691 (pMQ113), VTT-D-98681 (pMQ105)

and VTT-D-98682 (pMQ115). These strains have QM9414 as the host strain. VTT-D-99702 (pMQ113) has Rut-C30 as the host strain.

Example 4

Cultivation of the EGI-HFBI and EGIcore-HFBI protein producing and HFBI over-producing *Trichoderma* strains

EGI-HFBI and EGIcore-HFBI fusions were produced under the *cbh1* promoter in a 15-litre fermenter using *T. reesei* strains VTT-D-98693 (pMQ103) and VTT-D-98691 (pMQ113), respectively. Strains were grown 5 days on minimal medium (Penttilä et al. 1987) containing 4% Solka floc cellulose (James River Corporation, Berlin, NH) and 2 % spent grain (Primalco, Koskenkorva, Finland). EGIcore-HFBI was also produced in fermenter (15 l) using the Rut-C30 strain VTT-D-99702 (pMQ113) with 4 % lactose medium. To induce the production of EGI-HFBI and EGIcore-HFBI fusions under *Aspergillus gpdA* promoter, *T. reesei* strains VTT-D-98681 (pMQ105) and VTT-D-98682 (pMQ115) were cultivated in 15-litre fermenter. Strains were grown 3 to 5 days on minimal medium supplemented with 2% glucose, 0.2% Peptone, and 0.1% Yeast Extract, and with glucose feed to maintain the glucose concentration in the range of 1 to 3% throughout the cultivation. HFBI over-producing strain VTT-D-98692 (pEA10) was grown similarly in 15 l on glucose medium and the strain VTT-D-98492 (pQM121) over-producing HFBI under *cbh1* promoter was cultivated for 7 days in 15-litre fermentor on medium containing 4% Solka floc and 2% spent grain. The control cultivations with the host strains of the transformants, QM9414 (VTT-D-74075) and Rut-C30 (VTT-D-86271), were carried out on media containing i) Solka floc sellulose and either spent grain or whey, ii) lactose and iii) glucose similarly as described above.

When proper some *T. reesei* transformant strains and their host strains were also cultivated at 28°C in shake flasks for 5 to 6 days in 50 to 150 ml volume of *Trichoderma* minimal medium (Penttilä et el. 1987) suplemented with either 3% Solka floc cellulose and 1% spent grain or 3-4% glucose with glucose feeding.

Example 5

Standard separation assays and analysis

If not otherwise stated the standard ATPS and subsequent analyses and calculations were carried out as explained in this example.

In general whole fermentation broth, supernatand (biomass separated by centrifugation or filtration) or purified proteins in buffer were separated in 10 ml graduated tubes. First detergent was added into the tubes and the tubes were then filled to 10 mg with protein containing liquid. The amount of detergent in the tube was calculated in weight percent of detergents. After thorough mixing in an overhead shaker the separation took place by either gravity settling in a water bath at constant temperature or by centrifugation at constant temperature. The separation usually was performed at 30°C, the standard amount of detergent used was 2-5% (w/v). After separation the volume ratio was noted and samples were taken from the lighter and heavier phase for analysis.

Two-phase separations were analysed qualitatively by using SDS-PAGE gels followed by visualization of the fusion proteins with Coomassie brilliant blue R-250 (Sigma) or Western blotting. Polyclonal anti-HFBI antibody were used in Western analysis for detection of EGIcore-HFBI, EGI-HFBI and dCBD-HFBI proteins together with alkaline phosphatase conjugated anti-rabbit IgG (Bio-Rad). Alkaline phosphatase activity was detected colorimetrically with BCIP (5-bromo-4-chloro-3-indolyl-phosphate) used in conjunction with NBT (nitro blue tetrazolium) (Promega).

Contaminating endogenous EGI, CBHI and EGIII in the top phase was tested with appropriate antibodies. Acidic protease activity in the top and bottom phase was also tested using the SAP method (Food Chemicals Codex, p. 496-497, 1981), which is based on the 30 min enzymatic hydrolysis of a hemoglobin substrate. All reactions were performed at pH 4.7 and 40 °C. Unhydrolyzed substrate was precipitated with 14% TCA and removed by filtration. The released tyrosine and tryptophan was determined spectrophotometrically.
Total protein concentrations were determined by *Non-Interfering* Protein Assay™ (Geno Technology, Inc).

EGI activity was detected using 4-methylumbelliferyl-ß-D-cellobioside (MUC) (Sigma M 6018) as substrate (Van Tilbeurgh H. & Caeyssens M., 1985; Van Tilbeurgh *et.al.* ,1982) . EGI hydrolyses the ß-glycosidic bond and fluorogenic 4-methylumbelliferone is released, which can be measured using a fluorometer equipped with a 360 excitation filter and a 455 nm emission filter. CBHI also

hydrolyses the substrate and it was inhibited by addition of cellobiose (C-7252, Sigma). EGI containing liquid was added in an appropriate dilution to a buffer containing 50 mM sodium acetate buffer (pH 5) , 0.6 mM MUC and 4.6 mM cellobiose. The mixture was heated to 50°C. The reaction was stopped after ten minutes using 2% Na_2CO_3, pH 10. Purified CBHI was detected using the same assay as for EGI without the addition of the inhibitor cellobiose.

The partition coefficient K was defined as the ratio of the measured concentrations or activities in the top and bottom phase respectively.

The Yield Y was defined as follows:

$$Y_T = \frac{1}{1+\left[\frac{V_B}{V_T}\cdot\frac{1}{K}\right]}$$

where Y_T is the Yield of the top phase, V_B and V_T are the volumes of top and bottom phase respectively. The Yield of the bottom phase can be described accordingly.

The mass balances, e.g. recovery of all added protein, were always checked for completeness to ensure no artificially high Yield (e.g. due to possible inactivation of the protein in the bottom phase). The values were usually calculated based on total enzyme activity (EGI wt plus the EGI-fusion) and thus the values are underestimated for the separation of the fusion as demonstrated in Example 16.

Example 6

Small scale ATPS separation studies and gel analysis

EGI-HFBI and EGIcore-HFBI fusions produced under the *cbh1* promoter in a 15-litre fermenter on solka flock cellulose and spent grain medium as described in Example 4 using *T. reesei* strains VTT-D-98693 (pMQ103) and VTT-D-98691 (pMQ113), respectively, were separated in small scale ATPS as described above.

The phases from the two phase separations were analysed qualitatively by using SDS-PAGE gels followed by visualization of the fusion proteins with Coomassie brilliant blue or Western blotting. Coomassie stained SDS-PAGE (10%) is shown in Figure 8. In the lane containing the non-extracted culture filtrate three distinct closely migrating bands can be seen (the sample was diluted 1/10 with H_2O). The topmost band is CBHI, the band in the middle is EGIcore-HFBI fusion and the lower one endogenous EGI. In the samples separated in ATPS, only two bands (CBHI and EGI) are

seen in the sample from bottom phase and one band representing EGIcore-HFBI in the sample obtained from the top phase.

Western blotting with HFBI antibody showed thick bands for the top phase, whereas for the bottom phase there was only faint band demonstrating that the fusion is separating strongly into the detergent top phase. Figure 9 shows the separation of the EGIcore-HFBI fusion produced on cellulose media into the top phase. Contaminating endogenous EGI and EGIII in the top phase was tested with appropriate antibodies but no signal was detected.

Small amount of endogenous CBHI was found in the upper phase when CBHI antibody was used in Western blotting. EGI, EGIII and proteases were not found in the top phase. Further purification from the contaminating CBHI was observed when the top phase was re-extracted with 2 % detergent. The Figure 10 shows that the upper phase does not any more contain CBHI and pure fusion protein is recovered.

EGIcore-HFBI was also produced in fermenter (15 l) using the Rut-C30 strain VTT-D-99702 (pMQ113) with 4 % lactose medium. The separation in ATPS carried out in the standard manner gave essentially the same result as the separation from cellulose containing medium thus demonstrating that the purification can be carried out from several media relevant for large scale industrial use.

Acidic protease activity in the top was only 1/15 compared to the bottom phase (table below) demonstrating that acidic proteases remain in the bottom phase.

	A (275 nm)	HUT[3]/ml
Bottom phase[1]	0.146	41.6
Top phase[2]	0.009	2.6

[1]1/10 diluted bottom phase after separation of VTT-D-98691 culture filtrate with 2% detergent

[2]1/100 diluted bottom phase after separation of VTT-D-98691 culture filtrate with 2% detergent

[3]1 HUT = enzyme concentration, which in reaction conditions hydrolyses hemoglobin in 1 min so that the absorbance at 275 nm of the formed hydrolysate equals 1.10 µg tyrosine/ml 0.006 N HCl solution.

These results show that the fusion protein can be purified extremely efficiently and the resulting preparate is free of other proteins produced by the fungus including proteases.

Example 7

Recovery of the native EGI in ATPS after thrombin cleavage

EGI-HFBI protein produced by the strain VTT-D-98693 has a thrombin cleavage site (LVPRGS) designed in the linker region between the EGI CBD and HFBI, which would enable the recovery of the native EGI after thrombin cleavage. EGI-HFBI fusion protein was purified from the culture filtrate (100 ml) of strain VTT-D-98693 grown on 4% solka flock cellulose and 2 % spent grain as described in Example 4 using the 2-phase separation system (5 % detergent). After removal of the bottom phase the detergent phase was extracted by isobutanol. The resulting water phase (~19 ml) was divided in eppendorf tubes and the liquid was evaporated with speed vac. Remaining lyophilizate was diluted to 50 mM Tris-Cl (pH 8). To test the efficiency of thrombin cleavage, 9 units of thrombin (Sigma) was incubated > 24 h with 1 mg EGI-HFBI fusion protein in 36 °C at pH 8.0. Coomassie stained SDS-PAGE (10 %) was used for detection.

Only minor cleavage was observed in 48 h under these conditions (Figure 11), possibly due to steric hindrance by O-glycosylation in the linker.

Example 8

Separation of low concentrations of EGIcore-HFBI in ATPS

Detergent based aqueous two-phase systems were successfully applied using very low concentrations (diluted) of EGIcore-HFBI fusion protein produced with the *cbh1* promoter in *T. reesei* VTT-D-98691 (pMQ 113) from a 15 liter cultivation carried out on solka flock cellulose with spent grain as described in Example 4.

The original protein concentration of the supernatant was 7.0 mg/ml. This supernatant was diluted with de-ionised water by a factor of 100 and 1000, respectively. The fusion protein could be

separated using 2% (w/w) of the detergent C12-C18EO5 with partitioning coefficients higher than 5. This is shown in the table below together with the experiment with non-diluted supernatant. The partitioning coefficients were calculated based on activity measurements for total EGI (wild type and fusion protein together).

	EGIcore-HFBI non-diluted supernatant	dilution 1/100	dilution 1/1000
K	4.1	5.3	5.6
Y [%]	38	31	32

Example 9

Separation of EGIcore-HFBI from fungal biomass containing culture broths

EGIcore-HFBI from *T.reesei* strain VTT-D-98691 (pMQ 113-2) was cultivated (50 ml in 250 ml shake flasks) on solka flock cellulose with spent grain as described in Example 4. Directly after the cultivation, part of the whole broth was centrifuged at 3000 rpm for 30 min, supernatant was spilled out and the centrifuged mycelium was added to the supernatant to obtain artificial whole broths containing different amounts of biomass.

Using 5% of C12-C18EO5 in a 10 g experiment consisting of up to 50% wet biomass (weight of wet biomass divided by the sum of wet biomass and supernatant) could still be separated without any difficulties. The Yield remained in between 61 and 64 % and therefore it is not significantly different in comparison to the experiment carried out with supernatant only (without mycelium) (see table below). The total recovery of the fusion protein is even higher. This is most probably due to cell attached enzyme extracted in the ATPS increasing the total amount of EGI. The partitioning coefficient was calculated based on activity measurements for total EGI (wild type and fusion protein together).

	K	Y [%]
supernatant	5,5	62
25% of wet biomass in supernatant	7,3	66

40% of wet biomass in supernatant	6,4	61
50% of wet biomass in supernatant	7,6	64

Example 10

Separation of EGI-HFBI in ATPS

EGI-HFBI from *Trichoderma reesei* strain VTT-D-98693 (pMQ 103) from a 15 liter cultivation carried out on solka flock cellulose and spent grain as described in Example 4 was separated in a 10 g experiment using different amounts of C12-C18EO5. The partitioning coefficients are shown below. The partitioning coefficient was calculated based on activity measurements for total EG1 (wild type and fusion protein together), and as in previous examples the endogenous EGI is included in the partitioning coefficients.

Detergent [% w/w]	2	3	5	7
K	1,9	1,8	1,4	1,1

Example 11

Separation of EGIcore-HFBI in 50 ml

EGIcore-HFBI from *T. reesei* strain VTT-D-98691 (pMQ 113) cultivated in 15 liters using solka flock cellulose and spent grain as described in Example 4 was separated in Falcon tubes in a 50 g experiment using 5% of C12-C18EO5. A partition coefficient of 2.52 and a yield of 51 % could be obtained. The separation was performed at 30°C at 3000rpm for 30 minutes. The values are based on activity measurements for total EGI activity (wild type and fusion protein together) including endogenous EGI.

Example 12

Separation of EGIcore in ATPS using different detergents

EGIcore-HFBI from *T. reesei* strain VTT-D-98691 (pMQ 113) cultivated in 15 liters using solka flock cellulose with spent grain as described in Example 4 was separated in a 10 g experiment using 2% of detergent in each experiment. The detergents investigated in this example were C10 EO5, C12 EO5, C14 EO6 (each Nikko Chemicals, Japan), C12-C18 EO5 („Agrimul NRE 1205", Henkel, Germany), C12/14 5EO, C12/14 6EO (Clariant, Germany), C9/11 EO5.5 („Berrol 266", Akzo Nobel, Germany), Triton X-114 (Sigma, Germany). The partition coefficients and yields are listed below. The values are based on activity measurements for total EGI activity (wild type and fusion protein together) including endogenous EGI.

	K	Y(fusion) [%]
C10EO5	20	56
C12EO5	15	57
C12-C18EO5	14	66
C12/14 5EO	12	58
C12/14 6EO	14	62
C14EO6	11	54
C9/11 EO5,5	5	30
Triton X-114	0,16	53

Example 13

Separation of EGIcore-HFBI in ATPS from glucose grown cultures

EGIcore-HFBI was separated from cultivation of the *Trichoderma reesei* strain VTT-D-98682 (pMQ115) cultivated with glucose as described in Example 4. The supernatant was separated with 2% of the detergent C12-C18 EO5. The fusion protein could be partitioned with a K value of 2.4. In comparison, the K value for the native EGI is 0.3 when measured in a similar way for purified EGI.

Example 14

Separation of EGIcore-HFBI using different concentrations of detergent

EGIcore-HFBI from *T. reesei* VTT-D-98691 (pMQ 113) cultivated in 15 litres using solka flock with spent grain as described in Example 4 was separated in detergent based ATPS applying different amounts of the detergent C12-C18 EO5 on the cell free supernatant. The partitioning coefficients are shown in the table below. The corresponding gel electrophoresis and Western antibody-blots are shown in Figure 8 and Figure 9, respectively.

The values are based on activity measurements of total EGI activity.

Amount of detergent C12-C18 EO5	K	Yield (%)
1,0%	6,1	9
2,0%	4,1	38
3,5%	3,6	50
5,0%	2,9	55
7,5%	1,7	53
10,0%	1,1	58

Example 15

Re-extraction of the detergent phase

Detergent based ATPS was applied on EGIcore-HFBI fusion protein containing supernatant produced by the strain VTT-D-98691(pMQ 113) in a shake flask cultivation. The first extraction using C12-C18EO5 conducted under the standard conditions shows a partitioning coefficient of 16 and a yield of 72 % (wild type EGI measured together with fusion protein). The top phase was re-extracted in 10 mM sodium acetate buffer (pH 5) with 2% of detergent. A partitioning coefficient of 52 and a yield of 89 % could be obtained. In the re-extraction experiment of the bottom phase (2% of detergent), a small yield of 7.5% and a K of 0.8 of EGI activity were achieved. The partitioning coefficients were calculated based on activity measurements for total EGI (wild type and fusion protein together). Due to the wild type EGI present in the sample, the yield is at least 72% and the partitioning coefficient at least 16 in the first extraction. The SDS-PAGE results of both extraxtions are shown in Figure 10.

Separation step	K	Y[%]
2% detergent	16	72
reextraction top phase	52	89
reextraction bottom phase	0.8	7.5

Example 16

Separation of pure cellulases in ATPS

The effect of HFBI on partitioning and the final yield can further be demonstrated by comparing the extraction result of EGIcore-HFB fusion with extraction results obtained with purified wild type EGI and EGIcore. The fusion protein is partitioning more than 100 times better to the detergent phase (see table below).

The improvement on the partitioning of the purified fusion protein from the first extraction obtained in the re-extraction (see Example 15) can be explaned by the partitioning of the wild type EGI as demonstrated with purified wild type EGI in the table below. The wild type EGI lowers the partitioning coefficient in the first extraction (since EGI activity is measured from both top and botom phase), but the absence of it in the re-extraction increases the partitioning coefficient of the EGIcore-HFBI fusion. The purity can in addition be demonstrated by analysing the partitioning of pure CBHI, which is the major contaminating protein corresponding to about 50 % of all secreted *T.reesei* proteins. Pure CBHI has a partitioning coefficient of 0.5 and a yield of 3.6 and is therefore separated from the fusion protein.

Separation step	K	Y[%]
re-extraction of top phase	52	89
extraction of pure wild type EGI	0.3	2.2
extraction of pure EGI-core	0.3	2.3
extraction of pure CBHI	0.5	3.6

Using the definitions of K and Y and calculating mass balances, the ratio of the amount of EGI fusion protein to EGI wild type can be calculated. The „true" partition coefficients and Yields can be concluded from this. „True" means the values which would be detected if it would be possible to measure only the amount of EGI-fusion without measuring the amount for EGI wild type at the same time.

The fundament for the calculation is the re-extraction experiment. The re-extracted top phase is believed to be pure. An example of the measured values and the calculated "true" values based on this are shown in the table below for two cultivations of VTT-D-98691 (pMQ113) grown as described in Example 4.

cultivation vessel	cultivation substrate	K „with EG1 wt"	„true" K	Y [%] „with EG1 wt"	„true" Y [%]
15 liter fermenter	whey permeate	4	6	16	54
250 ml shake flask	cellulose	16	54	66	90

Example 17

HFBI and HFBII purification in ATPS

HFBI was produced by cultivating the *T. reesei* strain VTT-D-98692 (pEA10-103B) using glucose as substrate as described in Example 4. HFBI could be separated using 2% of the detergent C12-C18 EO5 with a partition coefficient higher than 20 under the standard conditions described.

HFBII was produced by cultivating the *T. reesei* strain VTT-D-74075 (QM9414) on whey spent grain as described in Example 4. HFBII could be separated using 2% of the detergent C12-C18 EO5, exceeding a partition coefficient of 10 under the standard conditions.

Both HFBI and HFBII hydrophobins are thus partitioning well to the upper phase in ATPS.

Example 18

Detergent based ATPS with additional NaCl

EGIcore-HFBI from which cultivation of T.reesei was separated in a 10 g experiment using 5% of C12-C18EO5. The partitioning coefficient of the supernatant was 3.5 with a volume ratio of 0.2. Using 1.1 % (w/v) NaCl the partitioning coefficient could be increased to 4.3 with a lower volume ratio of 0.14.

Example 19

Construction of an *E. coli* strain expressing a fusion protein HFBI-dCBD, containing hydrophobin I and double cellulose binding (CBD) domains

A 280 bp DNA fragment containing a modified *cbh2* linker region followed by the coding region of *hfb1* from Ser-23 to the STOP codon was amplified by PCR using the plasmid pARO1 (Nakari-Setälä et al. Eur J. Biochem. (1996) 235:248-255) as a template. The 5' primer was 5' TCT AGC AAG CTT **GGC TCT AGT TCT GGA ACC GCA CCA GGC GGC** AGC AAC GGC AAC GGC AAT GTT TGC (SEQ ID 14) and the 3' primer was 5' TCG TAC AAG CTT TCA AGC ACC GAC GGC GGT (SEQ ID 15). The sequences in bold in the 5' and 3' primers encode the modified CBHII linker (Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly) and a translational STOP, respectively, and the underlined AAGCTT in both primers is a HindIII site. The PCR fragment was purified from agarose gel, digested with HindIII and ligated to HindIII digested and SAP treated (Shrimp Alkaline Phosphatase, USB) pSP73 resulting in plasmid pTNS9.

For subsequent cloning of the modified CBHII linker-HFBI fragment to an *E. coli* expression vector, pTNS9 was digested with HindIII and the proper fragment was purified from agarose gel. This HindIII fragment was cloned to HindIII digested and SAP treated (Shrimp Alkaline Phosphatase, USB) B599 resulting in pTNS13 (Figure 12). The *E. coli* expression vector B599 is essentially the same as the one described by Linder et al. (J. Biol. Chem. (1996) 271:21268-21272) except that it is missing a STOP codon at the end of the protein coding sequence. It carries the coding sequence for a fusion protein containing CBHII CBD (41 N-terminal residues of CBHII) and CBHI CBD linked together via CBHI linker region (CBHI linker and CBD are the last 57 residues in CBHI). The expression and secretion of the fusion protein in B599 is under the cotrol of *tac* promoter and *pelB* signal sequence (Takkinen et al. Protein Eng. (1991) 4:837-841). pTNS13 expression vector thus

carries the coding region for a fusion protein of double CBD and HFBI linked in frame via the Gly-Ser-Ser-Ser-Gly-Thr-Ala-Pro-Gly-Gly peptide. This vector also contains the *amp* gene for selection of *E. coli* transformants. pTNS13 plasmid was transformed into *E. coli* strain RV308 (su-, Δ*lac*X74, *gal*ISII::OP308, *str*A) and this strain was used for production of the fusion protein.

Example 20

Separation of HFBI-dCDB molecules expressed in *E.coli* in ATPS

dCBD-HFBI was produced in *E. coli* strain RV 308 transformed with pTNS13 plasmid as described above. The inoculum of RV308/pTNS13 was grown to the exponential growth phase in LB medium containing ampicillin (0.1 g/l) and 1% glucose. Fermentation was carried out using mineral salt medium described by Pack *et al.* (1993) with glucose (feed) in 10 litre fermenter. During cultivation temperature was maintained at 28 °C and pH was controlled at 6.8 with NH_4OH. Cell growth was monitored by measuring OD_{600} and dry weight of biomass. The culture was induced with 50 µM (final concentration) IPTG (isopropyl-β-D-thiogalactopyranoside) at late-exponential growth phase (OD_{600}=50-60) to promote fusion protein production.

Two-phase separation analysis of dCBD-HFBI protein was performed using culture filtrate and 5% detergent in the total volume of 40 ml. Results from Western blotting showed that 2-phase separation with 5 % detergent in the standard way was highly specific also for the dCBD-HFBI fusion. Strong signal was observed in the sample from the detergent phase compared to the sample from the bottom phase as shown in Figure 13.

Example 21

Construction of yeast strains expressing HFBI-FloI fusion protein on the cell surface

For construction of a HFBI-FLO1 fusion protein expression casette, *hfb1* (SEQ ID 1) coding region (from Ser-23 to the STOP codon) was amplified with PCR using pARO1 (Nakari-Setälä et al., 1996) as a template and as a 5' primer TCT AGC TCT AGA AGC AAC GGC AAC GGC AAT GTT (SEQ ID 16) and as a 3' primer TGC TAG TCG ACC TGC TAG CAG CAC CGA CGG CGG TCT G (SEQ ID 17). The underlined sequences in the 5' and 3' primers are XbaI and NheI sites, respectively. The 0.225 bp PCR fragment was purified from agarose gel and ligated to pGEM-T vector (Promega) resulting in pTNS10. The *hfb1* fragment was released from pTNS10 with XbaI and NheI and ligated to pTNS15 cut with the same restriction enzymes. Plasmid pTNS15 (Figure 14) is essentially the same as plasmid pBR-ADH1-FLO1L by Watari et al. 1994 except that a NheI site in the pBR322 backbone has been replaced by a BglII site and a unique XbaI site is introduced by linker cloning in the unique AocI site preceding the putative signal sequence cleavage site. The

resulting plasmid pTNS18 (Figure 15) contains the complete expression casette for HFBI-FLO1 fusion protein in which HFBI substitutes the putative lectin domain from Ser-26 to Ser-319 in the yeast flocculin FLO1 (SEQ ID 18).

In the next step, yeast expression vector for production of HFBI-FLO1 fusion protein was constructed. The expression vector used as a backbone in the construct is pYES2 (Invitrogen) (SEQ ID 19) which is a high-copy episomal vector designed for inducible expression of recombinant proteins in *S. cerevisiae*. It carries *GAL1* promoter and *CYC1* terminator sequences which regulate transcription, and 2μ origin of replication and *URA3* gene for maintenance and selection in the host strain. The plasmid pTNS18 was digested with HindIII and the released 3.95 kb fragment containing the expression casette for HFBI-FLO1 was purified from agarose gel and ligated to pYES2 digested with HindIII. This ligation mixture was concentrated by standard ethanol precipitation. The ligation mixture should contain besides unligated fragments and uncorrect ligation products also molecules where the vector and insert are correctly ligated with each other to result in plasmid pTNS23 which carries the expression casette for HFBI-FLO1 operably linked to *GAL1* and *CYC1* terminator sequences.

The above ligation mixture was transformed using the LiAc method of Gietz et al. 1992 into a laboratory *S. cerevisiae* strain H452 (wild type W303-1A; Thomas and Rothstein, 1989). Transformant colonies able to grow on SC-URA plates were picked and streaked on selective plates. Nitrocellulose replicas were taken from the plates and treated for colony hybridization according to Sherman et al. 1983. To find those yeast colonies containing the pTNS23 plasmid, replicas were hybridized with digoxigenin labelled *hfb1* coding fragment after which an immunological detection was performed all according to the manufacturer (Boehringer Mannheim). Plasmids were recovered from several yeast colonies giving positivie hybridization signal by isolating total DNA and using this in electroporation of *E. coli*. Restriction mapping and sequencing were carried out to confirm that the pTNS23 plasmid in the yeast transformants was correct. One of the transformants carrying plasmid pTNS23 was chosen for further studies and was designated VTT-C-99315. The control strain for it is yeast strain H2155 which carries the plasmid pYES2 in H452 background.

Example 22

Separation of yeast cells expressing HFBI-FloI fusion protein in ATPS

The *Saccharomyces cerevisiae* strain VTT-C-99315 (vector pTNS23) and its control strain H2155 (vector pYES2) were cultivated on synthetic complete medium lacking uracil (SC-URA) (Sherman, 1991) with 2% galactose as the carbon source to give an A_{600} of approximately 4. Approximately 6.3 x 10^8 cells in their culture medium were taken to ATPS using 7% (w/v) C12-18EO5 detergent (Agrimul NRE from Henkel) in a total volume of 5 ml. ATPS was carried out using strandard protocol. After phase separation by gravity settling, the top detergent phase was clearly turbid in the case of the strain VTT-C-99315 in contrast to the control strain whose detergent phase was clear (Figure 16). Samples were taken from the top phases and dilution series from 10^{-1} to 10^{-5} were prepared in 0.9% NaCl and plated on YPD plates. After incubation at 30°C the amount of yeast colonies were calculated showing at least 70 times more yeast colonies of the strain VTT-C-99318 on YPD plates compared to the control strain. This clearly manifests that also in a system over-loaded with cells, separation to the detergent phase of cells expressing a hydrophobin on cell suface occurs.

References

P.A. Albertsson, Partition of Cell Particles and Macromolecules, John Wiley & Sons, New York, 1986;

H. Walter, D.E. Brooks, D. Fisher, Partitioning in Aqueous Two-Phase Systems, Academic Press, Inc., Orlando, 1985;

M.-R. Kula, Bioseparation, 1, 181-189, (1990))

G.C. Terstappen, R. A. Ramelmeier, M.-R. Kula, *J. Biotechnol.*, 28, 263-275, (1993)

Johansson, H.-O., Karlström, G., Tjerneld, F., Haynes, A. C., Journal of Chromatography B, 711 (1998) 3-17

Ramagauri Bhikhabhai, Gunnar Johansson and Göran Petterson (1984). Isolation of cellulolytic enzymes from *Trichoderma reesei* QM 9414. J.Appl.Biochem. 6, 336-345.

Jaakko Pere, Matti Siika-aho, Johanna Buchert and Liisa Viikari (1995). Effect of purified *Trichoderma reesei* cellulases on the fiber properties of kraft pulp. Tappi Journal 78, 71-78..

Suominen,P., Mäntylä,A., Saarelainen,R., Paloheimo,M., Fagerström,R., Parkkinen,E. and Nevalainen,H. (1992). Genetic engineering of *Trichoderma reesei* to produce suitable enzyme combinations for applications in the pulp and paper industry. In: M.Kuwahara and M.Shimada (eds.) Biotechnology in Pulp and Paper Industry p.439-445. (Proc. 5th Int.Conf.Biotechn. Pulp Pap.Ind., Kyoto, Japan, 1992

Berggren K., Veide A., Nygren P.-A., Tjerneld F. (1999). Genetic engineering of protein-peptide fusions for control in thermoseparating aqueous two-phase systems Biotechnol. Bioeng. 62, 135-144.

Hassinen C., Köhler K., Veide A. (1994) Polyethylene glycol - potassiuim phosphate aqueous two-phase systems: A study using model proteins. Biotechnol. Bioeng. 44: 626-635

Köhler K., Ljungquist C., Kondo A., Veide A., and Nilsson B. (1991) Engineering proteins to enhance their partition coefficient in aqueous two-phase systems. Bio/Technology 9: 642-646.

Wessels, 1994;

Wösten and Wessels, 1997;

Kershaw and Talbot, 1998

Lora *et al.* 1994;

Lora *et al.* 1995;

Arntz and Tudzynski, 1997

Temple et al. 1997.

Wösten *et al.*, 1993;

Wösten *et al.*, 1994b

Russo *et al.*, 1982;

Carpenter *et al.*, 1992

Wösten *et al.*, 1996 (Ustilago)

Bender W, Koglin B. Mechanische Trennung von Bioprodukten. Chem.-Ing.-Tech., 58, 565-577 (1986).

Bordier C. Phase Separation of Integral Membrane Proteins in Triton X-l 14 Solution. J. Biol. Chem., 25,1604-1607 (1981).

Heusch R, Kopp F. Structures in aqueous solutions of nonionic tensides. Progress in Col-
loid & Polymer Science, 77, 77-85 (1988).

Hustedt H, Kroner K, Kula MR. Applications of Phase Partitioning in Biotechnology. In: Walter H, Brooks DE; Fisher D (eds.), Partitioning in Aqueous Two-Phase Systems: Theory, Methods, Uses and Applications to Biotechnology, pp. 529-587 (1985).

Kula MR. Extraction and Purification of Enzymes Using Aqueous Two-Phase Systems. In: Wingard L; Goldstein, L (eds.), Applied Biochemistry and Bioengineering. New York, Academic Press, pp. 71-95 (1979).

Kula MR, Kroner KH, Hustedt H, SchüKe H. Technical aspects of extractive enzyme purification. Ann. N.Y. Acad. Sci., 369, 341-354 (1981).

Kula MR Kroner KH, Hustedt, H. Purification of enzymes by liquid-liquid extraction. In

Fiechter A (ed.), Advances in Biochemical Engineering, Springer Verlag, Berlin, pp. 74-1 1 8 (1982).

Kula MR. Liquid-Liquid Extraction of Biopolymers. In: Humphrey A; Cooney CL (eds.), Comprehensive Biotechnology, Vol. 2. Pergamon Press, New York, pp. 451-471 (1985).

Minuth T, Thömrnes J, Kula M-R Extraction of cholesterol oxidase from *Nocardia rhodochrous* using a nonionic surfactant-based aqueous two-phase system. J. Biotechnol., 38,151-164 (1995).

Minuth T, Thömmes J, Kula M-R. A closed concept for purification of the membrane-bound Cholesterol-Oxidase from Nocardia rhodochrous by surfactant-based cloud-point extraction, organic solvent extraction and anion-exchange chromatography. Biotechnol. Appl. Biochem, 23,107-116 (1996).

Tjerneld F, Bemer S, Cajarville A, Johansson G. New aqueous two-phase system based on hydroxypropyl starch useful in enzyme purification. Enz. Microb. Technol., 8, 417-423 (1986).

Veide A, Lindbäck T, Enfors S. Continuous extraction of ß-D-galactosidase from *Escherichia coli* in an aqueous two-phase system: effects of biomass concentration on partitioning and mass transfer. Enz. Microb. Technol., 6, 325-330 (1984).

Gietz, D., St. Jean, A., Woods, R.A. and Schiestl, R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. Nucl. Acid. Res. 20: 1425.

Hynes, M.J., Corrick, C.M. and King, J.A. (1983) Isolation of genomic clones containing the *amdS* gene of *Aspergillus nidulans* and their use in the analysis of structural and regulatory mutations. Mol. Cell Biol. 3:1430-1439.

Linder, M., Salovuori, I., Ruohonen, L. and Teeri, T.T. (1996) Characterization of a double cellulose-binding domain. J. Biol. Chem. 271:21268-21272

Mach et al. 1994

Nakari, T., Onnela, M.-L., Ilmén, M., Nevalainen, K.M.H. and Penttilä, M. (1994) Fungal promoters active in the presence of glucose. International Patent Appl. WO 94/04673.

Nakari-Setälä et al. Eur. J. Biochem. (1996) 235:248-255

Pack, P., Kujau, M., Schroeckh, V., Knupfer, U., Wenderoth, R., Riesenberg, D. and Pluckthun, A. (1993) Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of *Escherichia coli*. Biotechnology 11: 1271-1277

Penttilä, M, Nevalainen, H., Rättö, M., Salminen, E. and Knowles , J. (1987) A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei* Gene 61:155-164

Sherman, F., Fink, G.R. and Hicks, J.B. (1983) Methods in Yeast Genetics. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Takkinen, K., Laukkanen, M.-L., Sizmann, D., Alfthan, K., Immonen, T., Vanne, L., Kaartinen, M., Knowles, J.K.C. and Teeri, T.T. (1991) An active single-chain antibody containing a cellulase linker domain is secreted by *Escherichia coli*. Protein Eng. 4:837-841

Thomas, B.J. and Rothstein, R. (1989) Elevated recombination rates in transcriptionally active DNA. Cell 56: 619-630.

Tilburn, J., Scazzocchio, C., Taylor, G.G., Zabicky-Zissman, J.H., Lockington, R.A. and Davies, R.W. (1983) Transformation by integration in *Aspergillus nidulans* Gene 26:205-221

Van Tilbeurgh H., Caeyssens M., (1985) „Detection and differentiation of cellulase components using low molecular mass fluorogenic substrates", FEBS letters 187, 283-288

Van Tilbeurgh H., Caeyssens M, De Bruyne C. K., (1982) „The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes", FEBS letters 194, 152-156

Watari, J., Nomura, M., Sahara, H. and Koshino, S. (1994) Construction of flocculent brewer's yeast by chromosomal integration of the yeast flocculation gene *FLO1*. J. Inst. Brew. 100: 73-77.

Sequences

SEQ ID 1: Coding sequence of *hfb1* , 428 bp, introns underlined.

ATGAAGTT CTTCGCCATC GCCGCTCTCT TTGCCGCCGC

TGCCGTTGCC CAGCCTCTCG AGGACCGCAG CAACGGCAAC GGCAATGTTT GCCCTCCCGG

CCTCTTCAGC AACCCCCAGT GCTGTGCCAC CCAAGTCCTT GGCCTCATCG GCCTTGACTG

CAAAGTCCGT AAGTTGAGCC ATAACATAAG AATCCTCTTG ACGGAAATAT GCCTTCTCAC

TCCTTTACCC CTGAACAGCC TCCCAGAACG TTTACGACGG CACCGACTTC CGCAACGTCT

GCGCCAAAAC CGGCGCCCAG CCTCTCTGCT GCGTGGCCCC CGTTGTAAGT TGATGCCCCA

GCTCAAGCTC CAGTCTTTGG CAAACCCATT CTGACACCCA GACTGCAGGC CGGCCAGGCT

CTTCTGTGCC AGACCGCCGT CGGTGCTTGA

SEQ ID 2

TCG GG C ACT ACG TG C CAG TAT AGC AAC GAC TAC TAC TCG CAA TGC ***CTT GTT CCG CGT GGC TCT*** AGT TCT GGA ACC GCA

SEQ ID 3

TCG TAC GGA TCC TCA AGC ACC GAC GGC GGT

SEQ ID 4

ACT ACA CGG AG G AGC TC G ACG ACT TCG AGC AGC CCG AGC TGC ACG CAG AGC AAC GGC AAC GGC

SEQ ID 5: *T. reesei cbh1* promoter, 2211 bp

GAATTCTCAC GGTGAATGTA GGCCTTTTGT AGGGTAGGAA TTGTCACTCA AGCACCCCCA

ACCTCCATTA CGCCTCCCCC ATAGAGTTCC CAATCAGTGA GTCATGGCAC TGTTCTCAAA

TAGATTGGGG AGAAGTTGAC TTCCGCCCAG AGCTGAAGGT CGCACAACCG CATGATATAG

GGTCGGCAAC GGCAAAAAAG CACGTGGCTC ACCGAAAAGC AAGATGTTTG CGATCTAACA

TCCAGGAACC TGGATACATC CATCATCACG CACGACCACT TTGATCTGCT GGTAAACTCG

TATTCGCCCT AAACCGAAGT GCGTGGTAAA TCTACACGTG GGCCCCTTTC GGTATACTGC

GTGTGTCTTC TCTAGGTGCA TTCTTTCCTT CCTCTAGTGT TGAATTGTTT GTGTTGGGAG

TCCGAGCTGT AACTACCTCT GAATCTCTGG AGAATGGTGG ACTAACGACT ACCGTGCACC

TGCATCATGT ATATAATAGT GATCCTGAGA AGGGGGGTTT GGAGCAATGT GGGACTTTGA

TGGTCATCAA ACAAAGAACG AAGACGCCTC TTTTGCAAAG TTTTGTTTCG GCTACGGTGA

AGAACTGGAT ACTTGTTGTG TCTTCTGTGT ATTTTTGTGG CAACAAGAGG CCAGAGACAA

TCTATTCAAA CACCAAGCTT GCTCTTTTGA GCTACAAGAA CCTGTGGGGT ATATATCTAG

AGTTGTGAAG TCGGTAATCC CGCTGTATAG TAATACGAGT CGCATCTAAA TACTCCGAAG

CTGCTGCGAA CCCGGAGAAT CGAGATGTGC TGGAAAGCTT CTAGCGAGCG GCTAAATTAG

CATGAAAGGC TATGAGAAAT TCTGGAGACG GCTTGTTGAA TCATGGCGTT CCATTCTTCG

ACAAGCAAAG CGTTCCGTCG CAGTAGCAGG CACTCATTCC CGAAAAAACT CGGAGATTCC

TAAGTAGCGA TGGAACCGGA ATAATATAAT AGGCAATACA TTGAGTTGCC TCGACGGTTG

CAATGCAGGG GTACTGAGCT TGGACATAAC TGTTCCGTAC CCCACCTCTT CTCAACCTTT

GGCGTTTCCC TGATTCAGCG TACCCGTACA AGTCGTAATC ACTATTAACC CAGACTGACC

GGACGTGTTT TGCCCTTCAT TTGGAGAAAT AATGTCATTG CGATGTGTAA TTTGCCTGCT

TGACCGACTG GGGCTGTTCG AAGCCCGAAT GTAGGATTGT TATCCGAACT CTGCTCGTAG

AGGCATGTTG TGAATCTGTG TCGGGCAGGA CACGCCTCGA AGGTTCACGG CAAGGGAAAC

CACCGATAGC AGTGTCTAGT AGCAACCTGT AAAGCCGCAA TGCAGCATCA CTGGAAAATA

CAAACCAATG GCTAAAAGTA CATAAGTTAA TGCCTAAAGA AGTCATATAC CAGCGGCTAA

TAATTGTACA ATCAAGTGGC TAAACGTACC GTAATTTGCC AACGCGTTGT GGGGTTGCAG

AAGCAACGGC AAAGCCCACT TCCCACGTTT GTTTCTTCAC TCAGTCCAAT CTCAGCTGGT

GATCCCCCAA TTGGGTCGCT TGTTTGTTCC GGTGAAGTGA AAGAAGACAG AGGTAAGAAT

GTCTGACTCG GAGCGTTTTG CATACAACCA AGGGCAGTGA TGGAAGACAG TGAAATGTTG

ACATTCAAGG AGTATTTAGC CAGGGATGCT TGAGTGTATC GTGTAAGGAG GTTTGTCTGC

CGATACGACG AATACTGTAT AGTCACTTCT GATGAAGTGG TCCATATTGA AATGTAAGTC

GGCACTGAAC AGGCAAAAGA TTGAGTTGAA ACTGCCTAAG ATCTCGGGCC CTCGGGCTTC

GGCTTTGGGT GTACATGTTT GTGCTCCGGG CAAATGCAAA GTGTGGTAGG ATCGACACAC

TGCTGCCTTT ACCAAGCAGC TGAGGGTATG TGATAGGCAA ATGTTCAGGG GCCACTGCAT

GGTTTCGAAT AGAAAGAGAA GCTTAGCCAA GAACAATAGC CGATAAAGAT AGCCTCATTA

AACGAAATGA GCTAGTAGGC AAAGTCAGCG AATGTGTATA TATAAAGGTT CGAGGTCCGT
GCCTCCCTCA TGCTCTCCCC ATCTACTCAT CAACTCAGAT CCTCCAGGAG ACTTGTACAC
CATCTTTTGA GGCACAGAAA CCCAATAGTC AACCGCGGAC TGCGCATCAT G

SEQ ID 6: *T. reesei egl1* cDNA, 1588 bp

CCCCCCTATC TTAGTCCTTC TTGTTGTCCC AAAATGGCGC CCTCAGTTAC ACTGCCGTTG
ACCACGGCCA TCCTGGCCAT TGCCCGGCTC GTCGCCGCCC AGCAACCGGG TACCAGCACC
CCCGAGGTCC ATCCCAAGTT GACAACCTAC AAGTGTACAA AGTCCGGGGG GTGCGTGGCC
CAGGACACCT CGGTGGTCCT TGACTGGAAC TACCGCTGGA TGCACGACGC AAACTACAAC
TCGTGCACCG TCAACGGCGG CGTCAACACC ACGCTCTGCC CTGACGAGGC GACCTGTGGC
AAGAACTGCT TCATCGAGGG CGTCGACTAC GCCGCCTCGG GCGTCACGAC CTCGGGCAGC
AGCCTCACCA TGAACCAGTA CATGCCCAGC AGCTCTGGCG GCTACAGCAG CGTCTCTCCT
CGGCTGTATC TCCTGGACTC TGACGGTGAG TACGTGATGC TGAAGCTCAA CGGCCAGGAG
CTGAGCTTCG ACGTCGACCT CTCTGCTCTG CCGTGTGGAG AGAACGGCTC GCTCTACCTG
TCTCAGATGG ACGAGAACGG GGGCGCCAAC CAGTATAACA CGGCCGGTGC CAACTACGGG
AGCGGCTACT GCGATGCTCA GTGCCCCGTC CAGACATGGA GGAACGGCAC CCTCAACACT
AGCCACCAGG GCTTCTGCTG CAACGAGATG GATATCCTGG AGGGCAACTC GAGGGCGAAT
GCCTTGACCC CTCACTCTTG CACGGCCACG GCCTGCGACT CTGCCGGTTG CGGCTTCAAC
CCCTATGGCA GCGGCTACAA AAGCTACTAC GGCCCCGGAG ATACCGTTGA CACCTCCAAG
ACCTTCACCA TCATCACCCA GTTCAACACG GACAACGGCT CGCCCTCGGG CAACCTTGTG
AGCATCACCC GCAAGTACCA GCAAAACGGC GTCGACATCC CCAGCGCCCA GCCCGGCGGC
GACACCATCT CGTCCTGCCC GTCCGCCTCA GCCTACGGCG GCCTCGCCAC CATGGGCAAG
GCCCTGAGCA GCGGCATGGT GCTCGTGTTC AGCATTTGGA ACGACAACAG CCAGTACATG
AACTGGCTCG ACAGCGGCAA CGCCGGCCCC TGCAGCAGCA CCGAGGGCAA CCCATCCAAC
ATCCTGGCCA ACAACCCCAA CACGCACGTC GTCTTCTCCA ACATCCGCTG GGGAGACATT
GGGTCTACTA CGAACTCGAC TGCGCCCCCG CCCCCGCCTG CGTCCAGCAC GACGTTTTCG
ACTACACGGA GGAGCTCGAC GACTTCGAGC AGCCCGAGCT GCACGCAGAC TCACTGGGGG
CAGTGCGGTG GCATTGGGTA CAGCGGGTGC AAGACGTGCA CGTCGGGCAC TACGTGCCAG
TATAGCAACG ACTACTACTC GCAATGCCTT TAGAGCGTTG ACTTGCCTCT GGTCTGTCCA
GACGGGGGCA CGATAGAATG CGGGCACGCA GGGAGCTCGT AGACATTGGG CTTAATATAT

AAGACATGCT ATGTTGTATC TACATTAGCA AATGACAAAC AAATGAAAAA GAACTTATCA

AGCAAAAAAA AAAAAAAAAA AAAAAAAA

SEQ ID 7: *T. reesei cbh1* terminator, 745 bp

GGACCTACCC AGTCTCACTA CGGCCAGTGC GGCGGTATTG GCTACAGCGG CCCCACGGTC

TGCGCCAGCG GCACAACTTG CCAGGTCCTG AACCCTTACT ACTCTCAGTG CCTGTAAAGC

TCCGTGCGAA AGCCTGACGC ACCGGTAGAT TCTTGGTGAG CCCGTATCAT GACGGCGGCG

GGAGCTACAT GGCCCCGGGT GATTTATTTT TTTTGTATCT ACTTCTGACC CTTTTCAAAT

ATACGGTCAA CTCATCTTTC ACTGGAGATG CGGCCTGCTT GGTATTGCGA TGTTGTCAGC

TTGGCAAATT GTGGCTTTCG AAAACACAAA ACGATTCCTT AGTAGCCATG CATTTTAAGA

TAACGGAATA GAAGAAAGAG GAAATTAAAA AAAAAAAAAA AACAAACATC CCGTTCATAA

CCCGTAGAAT CGCCGCTCTT CGTGTATCCC AGTACCACGT CAAAGGTATT CATGATCGTT

CAATGTTGAT ATTGTTCCGC CAGTATGGCT CCACCCCCAT CTCCGCGAAT CTCCTCTTCT

CGAACGCGGT AGTGGCTGCT GCCAATTGGT AATGACCATA GGGAGACAAA CAGCATAATA

GCAACAGTGG AAATTAGTGG CGCAATAATT GAGAACACAG TGAGACCATA GCTGGCGGCC

TGGAAAGCAC TGTTGGAGAC CAACTTGTCC GTTGCGAGGC CAACTTGCAT TGCTGTCAAG

ACGATGACAA CGTAGCCGAG GACCC

SEQ ID 8

TAA CCG CGG T

SEQ ID 9

CTA GAC CGC GGT TAA T

SEQ ID 10: *T. reesei gpd1* promoter

GTCGAC
ACGATATACA GGCGCGGCTG ATGATAATGA TGATCGAGCA TGACTTGATG
CTGTATGTGA CAATATTGAC TGCGAGGAAC CATCAGGTGT GTATGGATGG
AATCATTCTG TAACCACCAA GGTGCATGCA TCATAAGGAT TCTCCTCAGC
TCACCAACAA CGAACGATGG CCATGTTAGT GAAGGCACCG TGATGGCAAG
ATAGAACCAC TATTGCATCT GCGCTTCCCA CGCACAGTAC GTCAAGTAAC

GTCAAAGCCG CCCTCCCGTA ACCTCGCCCG TTGTTGCTCC CCCCGATTGC
CTCAATCACA TAGTACCTAC CTATGCATTA TGGGCGGCCT CAACCCACCC
CCCCAGATTG AGAGCTACCT TACATCAATA TGGCCAGCAC CTCTTCGGCG
ATACATACTC GCCACCCCAG CCGGCGCGAT TGTGTGTACT AGGTAGGCTC
GTACTATACC AGCAGGAGAG GTGCTGCTTG GCAATCGTGC TCAGCTGTTA
GGTTGTACTT GTATGGTACT TGTAAGGTGG TCATGCAGTT GCTAAGGTAC
CTAGGGAGGG ATTCAACGAG CCCTGCTTCC AATGTCCATC TGGATAGGAT
GGCGGCTGGC GGGGCCGAAG CTGGGAACTC GCCAACAGTC ATATGTAATA
GCTCAAGTTG ATGATACCGT TTTGCCAGAT TAGATGCGAG AAGCAGCATG
AATGTCGCTC ATCCGATGCC GCATCACCGT TGTGTCAGAA ACGACCAAGC
TAAGCAACTA AGGTACCTTA CCGTCCACTA TCTCAGGTAA CCAGGTACTA
CCAGCTACCC TACCTGCCGT GCCTACCTGC TTTAGTGTTA ATCTTTCCAC
CTCCCTCCTC AATCTTCTTT TCCCTCCTCT CCTCTTTTTT TTTTCTTCCT
CCTCTTCTTC TCCATAACCA TTCCTAACAA CATCGACATT CTCTCCTAAT
CACCAGCCTC GCAAATCCTC AGTTTGTATG TACGTACGTA CTACAATCAT
CACCACGATC GTCCGCCCGA CGATGCGGCT TCTGTTCGCC TGCCCCTCCT
CTCACTCGTG CCCTTGACGA GCTAGCCCCG CCAGGACTCT CCTGCGTCAC
CAATTTTTTT CCCTATTTAC CCCTCCTCCC TCTCTCCCTC TCGTTTCTTC
CTAACAAACA ACCACCACCA AAATCTCTTT GGAAGCTCAC GACTCACGCA
GCTCAATTC GCAGATACAA ATCTAGA

SEQ ID 11: *T. reesei gpd1* terminator

GGATCC CGAGCATT GTCTATGAAT GCAAACAAAA ATAGTAAATA AATAGTAATT

CTGGCCATGA CGAATAGAGC CAATCTGCTC CACTTGACTA TCTTGTGACT

GTATCGTATG TCGAACCCTT GACTGCCCAT TCAAACAATT GTAAAGGAAT

ATAGCTACAA GTTATGTCTC ACGTTTGCGT GCGAGCCCGT TTGTACGTTA

TTTTGAGAAA GCGTTGCCAT CACATGCTCA CAGTCACTTG GCTTACGATC

ATGTTTGCGA TCTTCGGTAA GAATACACAG AGTAACGATT ATCTCCATCG

CTTCTATGAT TAGGTACTCA GACAACACAT GGGAAACAAG ATAACCATCG

CATGCAAGGT CGATTCCAAT CATGATCTGG ACTGGGGTAT TCCATCTAAG

CCATAGTACC CTCGAGAGAA GGAATGGTAG GACCTCTCAG GCGTCCACCA

TCTGTGCTGC AAATCCAAGA AACCCCCCAA AAGCACCTAC CTATCTACCT

AGAGTAACTG CACGAGAAAA GAAAAGGAGC AGAAGAAGAA TGATCTCAAG

AGGCCGTGAA CGCAGAAACA CACTCCTCCC AACTTTTCAA GTTTTGAACA

AAAAAAGAAA GATGAGGACT AGAAGATGGA GTATTTCCTT CTTAGAGAGC

TCTCGGTGAG GTGACCTGTC AGGGTTTACC GCAAACCGTC GGTGGTTCTA

TCCAATTAAT CAAGTCCCGC GCCTCGCCTC TTCTCTCCTG TCCTTTCATA

GAATCCCGTC TCCTTGTTGC TTGATCGAAG CGGGGTTATC GACGCCACCA

AAGATCTTGT CTTGGTGACT TATCAATCCT TTGGTGATCA AACAGCCCCC

GAGTGATCAG ATCCGTAAAA GAAGAAGAAG AGTACGATTT AACCAGACCG

AGGAACAATA AAGCGAGTAA ATAACATCAA AATAAGAGTC TCGTTGAAAA

TTACTTGTTC CTCAATCAAT CCCAACCCCC CTAAAAGCCC TTCCCCCCAT

GGTATATCCC GGCAGTAGGA GAGAGATATT TCCACTACCG CTCACCACCA

AGTGAGGCT TGCCGAGAGA AGAGGATGAA TCAGAAGTGA CAACAACGGG

TTGAGCACAT GGGATATC GGCGCGCC

SEQ ID12: Sequence of plasmid pAN52-1, 5733 bp

1-2129 *Aspergillus nidulans gpdA* promoter

2130-2304 *A. nidulans gpdA*

2305-3071 *A. nidulans trpC* terminator

3072-5726 pUC18 from SalI to EcoRI

SQ Sequence 5733 BP; 1435 A; 1454 C; 1378 G; 1463 T; 3 other;

```
CAATTCCCTT GTATCTCTAC ACACAGGCTC AAATCAATAA GAAGAACGGT TCGTCTTTTT      60
CGTTTATATC TTGCATCGTC CCAAAGCTAT TGGCGGGATA TTCTGTTTGC AGTTGGCTGA     120
CTTGAAGTAA TCTCTGCAGA TCTTTCGACA CTGAAATACG TCGAGCCTGC TCCGCTTGGA     180
AGCGGCGAGG AGCCTCGTCC TGTCACAACT ACCAACATGG AGTACGATAA GGGCCAGTTC     240
CGCCAGCTCA TTAAGAGCCA GTTCATGGGC GTTGGCATGA TGGCCGTCAT GCATCTGTAC     300
TTCAAGTACA CCAACGCTCT TCTGATCCAG TCGATCATCC GCTGAAGGCG CTTTCGAATC     360
TGGTTAAGAT CCACGTCTTC GGGAAGCCAG CGACTGGTGA CCTCCAGCGT CCCTTTAAGG     420
CTGCCAACAG CTTTCTCAGC CAGGGCCAGC CCAAGACCGA CAAGGCCTCC CTCCAGAACG     480
CCGAGAAGAA CTGGAGGGGT GGTGTCAAGG AGGAGTAAGC TCCTTATTGA AGTCGGAGGA     540
CGGAGCGGTG TCAAGAGGAT ATTCTTCGAC TCTGTATTAT AGATAAGATG ATGAGGAATT     600
GGAGGTAGCA TAGCTTCATT TGGATTTGCT TTCCAGGCTG AGACTCTAGC TTGGAGCATA     660
GAGGGTCCTT TGGCTTTCAA TATTCTCAAG TATCTCGAGT TTGAACTTAT TCCCTGTGAA     720
CCTTTTATTC ACCAATGAGC ATTGGAATGA ACATGAATCT GAGGACTGCA ATCGCCATGA     780
GGTTTTCGAA ATACATCCGG ATGTCGAAGG CTTGGGGCAC CTGCGTTGGT TGAATTTAGA     840
ACGTGGCACT ATTGATCATC CGATAGCTCT GCAAAGGGCG TTGCACAATG CAAGTCAAAC     900
GTTGCTAGCA GTTCCAGGTG GAATGTTATG ATGAGCATTG TATTAAATCA GGAGATATAG     960
CATGATCTCT AGTTAGCTCA CCACAAAAGT CAGACGGCGT AACCAAAAGT CACACAACAC    1020
AAGCTGTAAG GATTTCGGCA CGGCTACGGA AGACGGAGAA GCCACCTTCA GTGGACTCGA    1080
GTACCATTTA ATTCTATTTG TGTTTGATCG AGACCTAATA CAGCCCCTAC AACGACCATC    1140
AAAGTCGTAT AGCTACCAGT GAGGAAGTGG ACTCAAATCG ACTTCAGCAA CATCTCCTGG    1200
ATAAACTTTA AGCCTAAACT ATACAGAATA AGATAGGTGG AGAGCTTATA CCGAGCTCCC    1260
AAATCTGTCC AGATCATGGT TGACCGGTGC CTGGATCTTC CTATAGAATC ATCCTTATTC    1320
GTTGACCTAG CTGATTCTGG AGTGACCCAG AGGGTCATGA CTTGAGCCTA AAATCCGCCG    1380
CCTCCACCAT TTGTAGAAAA ATGTGACGAA CTCGTGAGCT CTGTACAGTG ACCGGTGACT    1440
CTTTCTGGCA TGCGGAGAGA CGGACGGACG CAGAGAGAAG GGCTGAGTAA TAAGCCACTG    1500
GCCAGACAGC TCTGGCGGCT CTGAGGTGCA GTGGATGATT ATTAATCCGG GACCGGCCGC    1560
CCCTCCGCCC CGAAGTGGAA AGGCTGGTGT GCCCCTCGTT GACCAAGAAT CTATTGCATC    1620
```

```
ATCGGAGAAT ATGGAGCTTC ATCGAATCAC CGGCAGTAAG CGAAGGAGAA TGTGAAGCCA      1680
GGGGTGTATA GCCGTCGGCG AAATAGCATG CCATTAACCT AGGTACAGAA GTCCAATTGC      1740
TTCCGATCTG GTAAAAGATT CACGAGATAG TACCTTCTCC GAAGTAGGTA GAGCGAGTAC      1800
CCGGCGCGTA AGCTCCCTAA TTGGCCCATC CGGCATCTGT AGGGCGTCCA AATATCGTGC      1860
CTCTCCTGCT TTGCCCGGTG TATGAAACCG GAAAGGCCGC TCAGGAGCTG GCCAGCGGCG      1920
CAGACCGGGA ACACAAGCTG GCAGTCGACC CATCCGGTGC TCTGCACTCG ACCTGCTGAG      1980
GTCCCTCAGT CCCTGGTAGG CAGCTTTGCC CCGTCTGTCC GCCCGGTGTG TCGGCGGGGT      2040
TGACAAGGTC GTTGCGTCAG TCCAACATTT GTTGCCATAT TTTCCTGCTC TCCCCACCAG      2100
CTGCTCTTTT CTTTTCTCTT TCTTTTCCCA TCTTCAGTAT ATTCATCTTC CCATCCAAGA      2160
ACCTTTATTT CCCCTAAGTA AGTACTTTGC TACATCCATA CTCCATCCTT CCCATCCCTT      2220
ATTCCTTTGA ACCTTTCAGT TCGAGCTTTC CCACTTCATC GCAGCTTGAC TAACAGCTAC      2280
CCCGCTTGAG CAGACATCAC CATGGATCCA CTTAACGTTA CTGAAATCAT CAAACAGCTT      2340
GACGAATCTG GATATAAGAT CGTTGGTGTC GATGTCAGCT CCGGAGTTGA GACAAATGGT      2400
GTTCAGGATC TCGATAAGAT ACGTTCATTT GTCCAAGCAG CAAAGAGTGC CTTCTAGTGA      2460
TTTAATAGCT CCATGTCAAC AAGAATAAAA CGCGTTTTCG GGTTTACCTC TTCCAGATAC      2520
AGCTCATCTG CAATGCATTA ATGCATTGAC TGCAACCTAG TAACGCCTTN CAGGCTCCGG      2580
CGAAGAGAAG AATAGCTTAG CAGAGCTATT TTCATTTTCG GGAGACGAGA TCAAGCAGAT      2640
CAACGGTCGT CAAGAGACCT ACGAGACTGA GGAATCCGCT CTTGGCTCCA CGCGACTATA      2700
TATTTGTCTC TAATTGTACT TTGACATGCT CCTCTTCTTT ACTCTGATAG CTTGACTATG      2760
AAAATTCCGT CACCAGCNCC TGGGTTCGCA AAGATAATTG CATGTTTCTT CCTTGAACTC      2820
TCAAGCCTAC AGGACACACA TTCATCGTAG GTATAAACCT CGAAATCANT TCCTACTAAG      2880
ATGGTATACA ATAGTAACCA TGCATGGTTG CCTAGTGAAT GCTCCGTAAC ACCCAATACG      2940
CCGGCCGAAA CTTTTTTACA ACTCTCCTAT GAGTCGTTTA CCCAGAATGC ACAGGTACAC      3000
TTGTTTAGAG GTAATCCTTC TTTCTAGAAG TCCTCGTGTA CTGTGTAAGC GCCCACTCCA      3060
CATCTCCACT CGACCTGCAG GCATGCAAGC TTGGCACTGG CCGTCGTTTT ACAACGTCGT      3120
GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC CCCTTTCGCC      3180
AGCTGGCGTA ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG      3240
AATGGCGAAT GGCGCCTGAT GCGGTATTTT CTCCTTACGC ATCTGTGCGG TATTTCACAC      3300
CGCATATGGT GCACTCTCAG TACAATCTGC TCTGATGCCG CATAGTTAAG CCAGCCCCGA      3360
CACCCGCCAA CACCCGCTGA CGCGCCCTGA CGGGCTTGTC TGCTCCCGGC ATCCGCTTAC      3420
AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA GGTTTTCACC GTCATCACCG      3480
AAACGCGCGA GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA TGTCATGATA      3540
ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG AACCCCTATT      3600
TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA      3660
ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT      3720
ATTCCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA      3780
GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC      3840
AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT      3900
AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT      3960
CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC AGAAAAGCAT      4020
CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC      4080
ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG      4140
CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC      4200
ATACCAAACG ACGAGCGTGA CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA      4260
CTATTAACTG GCGAACTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG      4320
GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT      4380
GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT      4440
GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA      4500
CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC      4560
CAAGTTTACT CATATATACT TTAGATTGAT TTAAAACTTC ATTTTTAATT TAAAAGGATC      4620
TAGGTGAAGA TCCTTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC      4680
```

```
CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTTCTG      4740
CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG      4800
GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACTGGCT TCAGCAGAGC GCAGATACCA      4860
AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG      4920
CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG      4980
TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA      5040
ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC      5100
CTACAGCGTG AGCTATGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT      5160
CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC      5220
TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA      5280
TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCTT TTTACGGTTC      5340
CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC TGATTCTGTG      5400
GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG      5460
CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA TACGCAAACC GCCTCTCCCC      5520
GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC      5580
AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC      5640
TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA      5700
AACAGCTATG ACCATGATTA CGAATTGCGG CCG                                   5733
```

SEQ ID 13

GTC AA C CGC GG A CTG CGC ATC ATG AAG TTC TTC GCC ATC

SEQ ID 14

TCT AGC AAG CTT GGC TCT AGT TCT GGA ACC GCA CCA GGC GGC AGC AAC GGC AAC GGC AAT GTT TGC

SEQ ID 15

TCG TAC AAG CTT TCA AGC ACC GAC GGC GGT

SEQ ID 16

TCT AGC TCT AGA AGC AAC GGC AAC GGC AAT GTT

SEQ ID 17

TGC TAG TCG ACC TGC TAG CAG CAC CGA CGG CGG TCT G

SEQ ID 18: *S. cerevisiae FLO1* coding sequence, 4614 bp

ATGACAATGC CTCATCGCTA TATGTTTTTG GCAGTCTTTA CACTTCTGGC ACTAACTAGT

GTGGCCTCAG GAGCCACAGA GGCGTGCTTA CCAGCAGGCC AGAGGAAAAG TGGGATGAAT

ATAAATTTTT ACCAGTATTC ATTGAAAGAT TCCTCCACAT ATTCGAATGC AGCATATATG

GCTTATGGAT ATGCCTCAAA AACCAAACTA GGTTCTGTCG GAGGACAAAC TGATATCTCG

ATTGATTATA ATATTCCCTG TGTTAGTTCA TCAGGCACAT TTCCTTGTCC TCAAGAAGAT

TCCTATGGAA ACTGGGGATG CAAAGGAATG GGTGCTTGTT CTAATAGTCA AGGAATTGCA

TACTGGAGTA CTGATTTATT TGGTTTCTAT ACTACCCCAA CAAACGTAAC CCTAGAAATG

ACAGGTTATT TTTTACCACC ACAGACGGGT TCTTACACAT TCAAGTTTGC TACAGTTGAC

GACTCTGCAA TTCTATCAGT AGGTGGTGCA ACCGCGTTCA ACTGTTGTGC TCAACAGCAA

CCGCCGATCA CATCAACGAA CTTTACCATT GACGGTATCA AGCCATGGGG TGGAAGTTTG

CCACCTAATA TCGAAGGAAC CGTCTATATG TACGCTGGCT ACTATTATCC AATGAAGGTT

GTTTACTCGA ACGCTGTTTC TTGGGGTACA CTTCCAATTA GTGTGACACT TCCAGATGGT

ACCACTGTAA GTGATGACTT CGAAGGGTAC GTCTATTCCT TTGACGATGA CCTAAGTCAA

TCTAACTGTA CTGTCCCTGA CCCTTCAAAT TATGCTGTCA GTACCACTAC AACTACAACG

GAACCATGGA CCGGTACTTT CACTTCTACA TCTACTGAAA TGACCACCGT CACCGGTACC

AACGGCGTTC CAACTGACGA AACCGTCATT GTCATCAGAA CTCCAACAAC TGCTAGCACC

ATCATAACTA CAACTGAGCC ATGGAACAGC ACTTTTACCT CTACTTCTAC CGAATTGACC

ACAGTCACTG GCACCAATGG TGTACGAACT GACGAAACCA TCATTGTAAT CAGAACACCA

ACAACAGCCA CTACTGCCAT AACTACAACT GAGCCATGGA ACAGCACTTT TACCTCTACT

TCTACCGAAT TGACCACAGT CACCGGTACC AATGGTTTGC CAACTGATGA GACCATCATT

GTCATCAGAA CACCAACAAC AGCCACTACT GCCATGACTA CAACTCAGCC ATGGAACGAC

ACTTTTACCT CTACATCCAC TGAAATGACC ACCGTCACCG GTACCAACGG TTTGCCAACT

GATGAAACCA TCATTGTCAT CAGAACACCA ACAACAGCCA CTACTGCTAT GACTACAACT

CAGCCATGGA ACGACACTTT TACCTCTACA TCCACTGAAA TGACCACCGT CACCGGTACC

AACGGTTTGC CAACTGATGA AACCATCATT GTCATCAGAA CACCAACAAC AGCCACTACT

GCCATGACTA CAACTCAGCC ATGGAACGAC ACTTTTACCT CTACATCCAC TGAAATGACC

ACCGTCACCG GTACCAATGG TTTGCCAACT GATGAGACCA TCATTGTCAT CAGAACACCA

ACAACAGCCA CTACTGCCAT GACTACAACT CAGCCATGGA ACGACACTTT TACCTCTACA

TCCACTGAAA TGACCACCGT CACCGGTACC AACGGTTTGC CAACTGATGA AACCATCATT

GTCATCAGAA CACCAACAAC AGCCACTACT GCCATAACTA CAACTGAGCC ATGGAACAGC

ACTTTTACCT CTACTTCTAC CGAATTGACC ACAGTCACCG GTACCAATGG TTTGCCAACT

GATGAGACCA TCATTGTCAT CAGAACACCA ACAACAGCCA CTACTGCCAT GACTACAACT

CAGCCATGGA ACGACACTTT TACCTCTACA TCCACTGAAA TGACCACCGT CACCGGTACC

AACGGTTTGC CAACTGATGA AACCATCATT GTCATCAGAA CACCAACAAC AGCCACTACT

GCCATGACTA CAACTCAGCC ATGGAACGAC ACTTTTACCT CTACATCCAC TGAAATGACC

ACCGTCACCG GTACCAACGG TTTGCCAACT GATGAGACCA TCATTGTCAT CAGAACACCA

ACAACAGCCA CTACTGCCAT GACTACAACT CAGCCATGGA ACGACACTTT TACCTCTACA

TCCACTGAAA TGACCACCGT CACCGGTACC AACGGCGTTC CAACTGACGA AACCGTCATT

GTCATCAGAA CTCCAACTAG TGAAGGTCTA ATCAGCACCA CCACTGAACC ATGGACTGGT

ACTTTCACCT CTACATCCAC TGAGATGACC ACCGTCACCG GTACTAACGG TCAACCAACT

GACGAAACCG TGATTGTTAT CAGAACTCCA ACCAGTGAAG GTTTGGTTAC AACCACCACT

GAACCATGGA CTGGTACTTT TACTTCTACA TCTACTGAAA TGACCACCAT TACTGGAACC

AACGGCGTTC CAACTGACGA AACCGTCATT GTCATCAGAA CTCCAACCAG TGAAGGTCTA

ATCAGCACCA CCACTGAACC ATGGACTGGT ACTTTTACTT CTACATCTAC TGAAATGACC

ACCATTACTG GAACCAATGG TCAACCAACT GACGAAACCG TTATTGTTAT CAGAACTCCA

ACTAGTGAAG GTCTAATCAG CACCACCACT GAACCATGGA CTGGTACTTT CACTTCTACA

TCTACTGAAA TGACCACCGT CACCGGTACC AACGGCGTTC CAACTGACGA AACCGTCATT

GTCATCAGAA CTCCAACCAG TGAAGGTCTA ATCAGCACCA CCACTGAACC ATGGACTGGC

ACTTTCACTT CGACTTCCAC TGAGGTTACC ACCATCACTG GAACCAACGG TCAACCAACT

GACGAAACTG TGATTGTTAT CAGAACTCCA ACCAGTGAAG GTCTAATCAG CACCACCACT

GAACCATGGA CTGGTACTTT CACTTCTACA TCTGCTGAAA TGACCACCGT CACCGGTACT

AACGGTCAAC CAACTGACGA AACCGTGATT GTTATCAGAA CTCCAACCAG TGAAGGTTTG

GTTACAACCA CCACTGAACC ATGGACTGGT ACTTTTACTT CGACTTCCAC TGAAATGTCT

ACTGTCACTG GAACCAATGG CTTGCCAACT GATGAAACTG TCATTGTTGT CAAAACTCCA

ACTACTGCCA TCTCATCCAG TTTGTCATCA TCATCTTCAG GACAAATCAC CAGCTCTATC

ACGTCTTCGC GTCCAATTAT TACCCCATTC TATCCTAGCA ATGGAACTTC TGTGATTTCT

TCCTCAGTAA TTTCTTCCTC AGTCACTTCT TCTCTATTCA CTTCTTCTCC AGTCATTTCT

TCCTCAGTCA TTTCTTCTTC TACAACAACC TCCACTTCTA TATTTTCTGA ATCATCTAAA

TCATCCGTCA TTCCAACCAG TAGTTCCACC TCTGGTTCTT CTGAGAGCGA AACGAGTTCA

GCTGGTTCTG TCTCTTCTTC CTCTTTTATC TCTTCTGAAT CATCAAAATC TCCTACATAT

TCTTCTTCAT CATTACCACT TGTTACCAGT GCGACAACAA GCCAGGAAAC TGCTTCTTCA

TTACCACCTG CTACCACTAC AAAAACGAGC GAACAAACCA CTTTGGTTAC CGTGACATCC

TGCGAGTCTC ATGTGTGCAC TGAATCCATC TCCCCTGCGA TTGTTTCCAC AGCTACTGTT

ACTGTTAGCG GCGTCACAAC AGAGTATACC ACATGGTGCC CTATTTCTAC TACAGAGACA

ACAAAGCAAA CCAAAGGGAC AACAGAGCAA ACCACAGAAA CAACAAAACA AACCACGGTA

GTTACAATTT CTTCTTGTGA ATCTGACGTA TGCTCTAAGA CTGCTTCTCC AGCCATTGTA

TCTACAAGCA CTGCTACTAT TAACGGCGTT ACTACAGAAT ACACAACATG GTGTCCTATT

TCCACCACAG AATCGAGGCA ACAAACAACG CTAGTTACTG TTACTTCCTG CGAATCTGGT

GTGTGTTCCG AAACTGCTTC ACCTGCCATT GTTTCGACGG CCACGGCTAC TGTGAATGAT

GTTGTTACGG TCTATCCTAC ATGGAGGCCA CAGACTGCGA ATGAAGAGTC TGTCAGCTCT

AAAATGAACA GTGCTACCGG TGAGACAACA ACCAATACTT TAGCTGCTGA AACGACTACC

AATACTGTAG CTGCTGAGAC GATTACCAAT ACTGGAGCTG CTGAGACGAA AACAGTAGTC

ACCTCTTCGC TTTCAAGATC TAATCACGCT GAAACACAGA CGGCTTCCGC GACCGATGTG

ATTGGTCACA GCAGTAGTGT TGTTTCTGTA TCCGAAACTG GCAACACCAA GAGTCTAACA

AGTTCCGGGT TGAGTACTAT GTCGCAACAG CCTCGTAGCA CACCAGCAAG CAGCATGGTA

GGATATAGTA CAGCTTCTTT AGAAATTTCA ACGTATGCTG GCAGTGCCAA CAGCTTACTG

GCCGGTAGTG GTTTAAGTGT CTTCATTGCG TCCTTATTGC TGGCAATTAT TTAA

SEQ ID 19: Sequence of pYES2

Comments for pYES2:

5857 nucleotides

GAL1 promoter: bases 1-452

T7 promoter/priming site: bases 476-495

Multiple cloning site: bases 502-601

CYC1 transcription terminator: bases 609-857

pMB1 (pUC-derived) origin: bases 1039-1712

Ampicillin resistance gene: bases 1857-2717

URA3 gene: bases 2735-3842

2 micron origin: bases 3846-5317

f1 origin: bases 5385-5840

ACGGATTAGAAGCCGCCGAGCGGGTGACAGCCCTCCGAAGGAAGACTCTCCTCCGTGCGTCCTCGTCCTC
ACCGGTCGCGTTCCTGAAACGCAGATGTGCCTCGCGCCGCACTGCTCCGAACAATAAAGATTCTACAATA
CTAGCTTTTATGGTTATGAAGAGGAAAAATTGGCAGTAACCTGGCCCCACAAACCTTCAAATGAACGAAT
CAAATTAACAACCATAGGATGATAATGCGATTAGTTTTTTAGCCTTATTTCTGGGGTAATTAATCAGCGA
AGCGATGATTTTTGATCTATTAACAGATATATAAATGCAAAAACTGCATTAACCACTTTAACTAATACTT
TCAACATTTTCGGTTTGTATTACTTCTTATTCAAATGTAATAAAAGTATCAACAAAAAATTGTTAATATA
CCTCTATACTTTAACGTCAAGGAGAAAAAACCCCGGATCGGACTACTAGCAGCTGTAATACGACTCACTA

TAGGGAATATTAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCTGCA
GATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCGCATCATGTAATTAGTTATGTCAC
GCTTACATTCACGCCCTCCCCCCACATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTA
GGTCCCTATTTATTTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTT
TTTCTGTACAGACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCT
CGAAGGCTTTAATTTGCGGCCCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG
GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGC
TCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA
GGCCAGCAAAAGCCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
GTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC
AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGA
CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG
CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGA
TCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA
TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC
AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCA
GCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGC
GCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATTCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGGCATTG
CTACAGGCATCGTGGTGTCACTCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGA
AGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCAT
CCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC
GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAGTGTATCACATAGCAGAACTTTAAAAGTGCTCATC
ATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC
CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG
AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTT
CAATGGGTAATAACTGATATAATTAAATTGAAGCTCTAATTTGTGAGTTTAGTATACATGCATTTACTTA
TAATACAGTTTTTTAGTTTTGCTGGCCGCATCTTCTCAAATATGCTTCCCAGCCTGCTTTTCTGTAACGT
TCACCCTCTACCTTAGCATCCCTTCCCTTTGCAAATAGTCCTCTTCCAACAATAATAATGTCAGATCCTG
TAGAGACCACATCATCCACGGTTCTATACTGTTGACCCAATGCGTCTCCCTTGTCATCTAAACCCACACC
GGGTGTCATAATCAACCAATCGTAACCTTCATCTCTTCCACCCATGTCTCTTTGAGCAATAAAGCCGATA
ACAAAATCTTTGTCGCTCTTCGCAATGTCAACAGTACCCTTAGTATATTCTCCAGTAGATAGGGAGCCCT
TGCATGACAATTCTGCTAACATCAAAAGGCCTCTAGGTTCCTTTGTTACTTCTTCTGCCGCCTGCTTCAA
ACCGCTAACAATACCTGGGCCCACCACACCGTGTGCATTCGTAATGTCTGCCCATTCTGCTATTCTGTAT

ACACCCGCAGAGTACTGCAATTTGACTGTATTACCAATGTCAGCAAATTTTCTGTCTTCGAAGAGTAAAA
AATTGTACTTGGCGGATAATGCCTTTAGCGGCTTAACTGTGCCCTCCATGGAAAAATCAGTCAAGATATC
CACATGTGTTTTTAGTAAACAAATTTTGGGACCTAATGCTTCAACTAACTCCAGTAATTCCTTGGTGGTA
CGAACATCCAATGAAGCACACAAGTTTGTTTGCTTTTCGTGCATGATATTAAATAGCTTGGCAGCAACAG
GACTAGGATGAGTAGCAGCACGTTCCTTATATGTAGCTTTCGACATGATTTATCTTCGTTTCCTGCAGGT
TTTTGTTCTGTGCAGTTGGGTTAAGAATACTGGGCAATTTCATGTTTCTTCAACACTACATATGCGTATA
TATACCAATCTAAGTCTGTGCTCCTTCCTTCGTTCTTCCTTCTGTTCGGAGATTACCGAATCAAAAAAAT
TTCAAAGAAACCGAAATCAAAAAAAAGAATAAAAAAAAAATGATGAATTGAATTGAAAAGCTAGCTTATC
GATGATAAGCTGTCAAAGATGAGAATTAATTCCACGGACTATAGACTATACTAGATACTCCGTCTACTGT
ACGATACACTTCCGCTCAGGTCCTTGTCCTTTAACGAGGCCTTACCACTCTTTTGTTACTCTATTGATCC
AGCTCAGCAAAGGCAGTGTGATCTAAGATTCTATCTTCGCGATGTAGTAAAACTAGCTAGACCGAGAAAG
AGACTAGAAATGCAAAAGGCACTTCTACAATGGCTGCCATCATTATTATCCGATGTGACGCTGCAGCTTC
TCAATGATATTCGAATACGCTTTGAGGAGATACAGCCTAATATCCGACAAACTGTTTTACAGATTTACGA
TCGTACTTGTTACCCATCATTGAATTTTGAACATCCGAACCTGGGAGTTTTCCCTGAAACAGATAGTATA
TTTGAACCTGTATAATAATATATAGTCTAGCGCTTTACGGAAGACAATGTATGTATTTCGGTTCCTGGAG
AAACTATTGCATCTATTGCATAGGTAATCTTGCACGTCGCATCCCCGGTTCATTTTCTGCGTTTCCATCT
TGCACTTCAATAGCATATCTTTGTTAACGAAGCATCTGTGCTTCATTTTGTAGAACAAAAATGCAACGCG
AGAGCGCTAATTTTTCAAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGC
TATTTTACCAACGAAGAATCTGTGCTTCATTTTTGTAAAACAAAAATGCAACGCGACGAGAGCGCTAATT
TTTCAAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTTACCAAC
AAAGAATCTATACTTCTTTTTTGTTCTACAAAAATGCATCCCGAGAGCGCTATTTTTCTAACAAAGCATC
TTAGATTACTTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTC
CGTTAAGGTTAGAAGAAGGCTACTTTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACTTCC
CGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCGATTATATTCT
ATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATTCTTCATTGGTCAGAA
AATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAATGTTTACATTTTCGTATTGT
TTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTCTAAAGAGTAATACTAGAGATAAACAT
AAAAAATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAGGGA
TATAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATGTTTGTGGAAGCGGTATTCGCAATGGGA
AGCTCCACCCCGGTTGATAATCAGAAAAGCCCCAAAAACAGGAAGATTGTATAAGCAAATATTTAAATTG
TAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACGAATAGCC
CGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTCC
AACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAAGGGTCTATCAGGGCGATG
GCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCAGTAAATCGGAA
GGGTAAACGGATGCCCCCATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAG
AAAGCGAAAGGAGCGGGGGCTAGGGCGGTGGGAAGTGTAGGGGTCACGCTGGGCGTAACCACCACACCCG
CCGCGCTTAATGGGGCGCTACAGGGCGCGTGGGGATGATCCACTAGT

Claims

1. A process for partitioning of molecules in aqueous two-phase systems (ATPS), comprising the steps of

a) constructing a fusion molecule by combining a molecule of interest to a targeting protein having the ability to carry said molecule of interest into one of the phases, and

b) subjecting said fusion molecule to an ATPS separation.

2. The process according to claim 1, wherein the targeting protein is a hydrophobic protein.

3. The process according to claim 1, wherein the targeting protein is selected from a group consisting of amphipathic proteins and proteins which form amphipathic aggregates.

4. The process according to claim 1, wherein the targeting protein is a hydrophobin-like protein.

5. The process according to claim 4, wherein the hydrophobin-like protein is a class 2 hydrophobin.

6. The process according to claim 5, wherein the hydrophobin is a *Trichoderma* hydrophobin.

7. The process according to claim 6, wherein the *Trichoderma* hydrophobin is HFBI, HFBII or SRHI.

8. A process for partitioning of particles, wherein the particles contain a targeting protein as defined in any one of claims 2 to 7 or a part thereof on their surface, the process comprising the step of

- subjecting the particles to ATPS separation.

9. The process according to claim 8, wherein the particles are cells.

10. The process according to claim 9, wherein the cells are yeast cells.

11. The process according to claim 8, wherein the particles are spores.

12. The process according to any one of claims 8 to 11, wherein the targeting protein is fused to a molecule which brings the targeting protein onto the surface of the particle.

13. The process according to any one of claims 1 to 12, wherein the aqueous two-phase system is selected from the group consisting of PEG/salt, PEG/Dextran and PEG/starch systems or derivatives thereof, detergent-based aqueous two-phase systems and thermoseparating polymer systems.

14. The process according to claim 13, wherein the detergent-based ATPS comprises a detergent which is selected from the group consisting of nonionic or zwitterionic detergents.

15. The process according to claim 13, wherein the thermoseparating polymer system comprises a polymer which is selected from the group consisting of polyethylene-polypropylene copolymers.

16. The process according to any one of claims 1-15, wherein the molecule of interest or the particle is separated from a suspension containing cells or cell extracts.

17. A fusion molecule, comprising a hydrophobin-like protein as defined in any one of claims 4 to 7 fused to a molecule of interest.

18. A fusion molecule according to claim 17, wherein the molecule of interest is a cell-bound protein or a part thereof.

19. The fusion molecule according to claim 17, wherein the molecule of interest is an extracellular protein or a part thereof.

20. The fusion molecule according to claim 19, wherein the extracellular protein is an extracellular protein of *Trichoderma*, selected from the group consisting of cellulases, hemicellulases and proteases.

21. The fusion molecule according to claim 17, wherein the molecule of interest is an antibody molecule or a part thereof.

22. The process according to claim 1, wherein the targeting protein is fused to the molecule of interest according to any one of the claims 18 to 21.

23. A recombinant organism producing a fusion molecule according to any one of claims 17 to 21.

24. A recombinant DNA molecule, comprising a DNA molecule encoding a fusion molecule according to any one of claims 17 to 21.

25. A process for producing a targeting protein as defined in any one of claims 4 to 7, or a fusion molecule according to any one of the claims 17 to 21 with recombinant organisms, the process comprising the steps of

a) transforming the recombinant organism with DNA molecules enabling expression of such molecules, and

b) recovering such molecules from the culture of the recombinant organism.

26. A process for separating hydrophobin-like molecules in aqueous two-phase systems, the process comprising the steps of

a) mixing solutions containing said hydrophobin-like molecule with the phase forming chemicals, and

b) carrying out ATPS separation,

wherein the aqueous two-phase system is as defined in any one of claims 13 to 15.

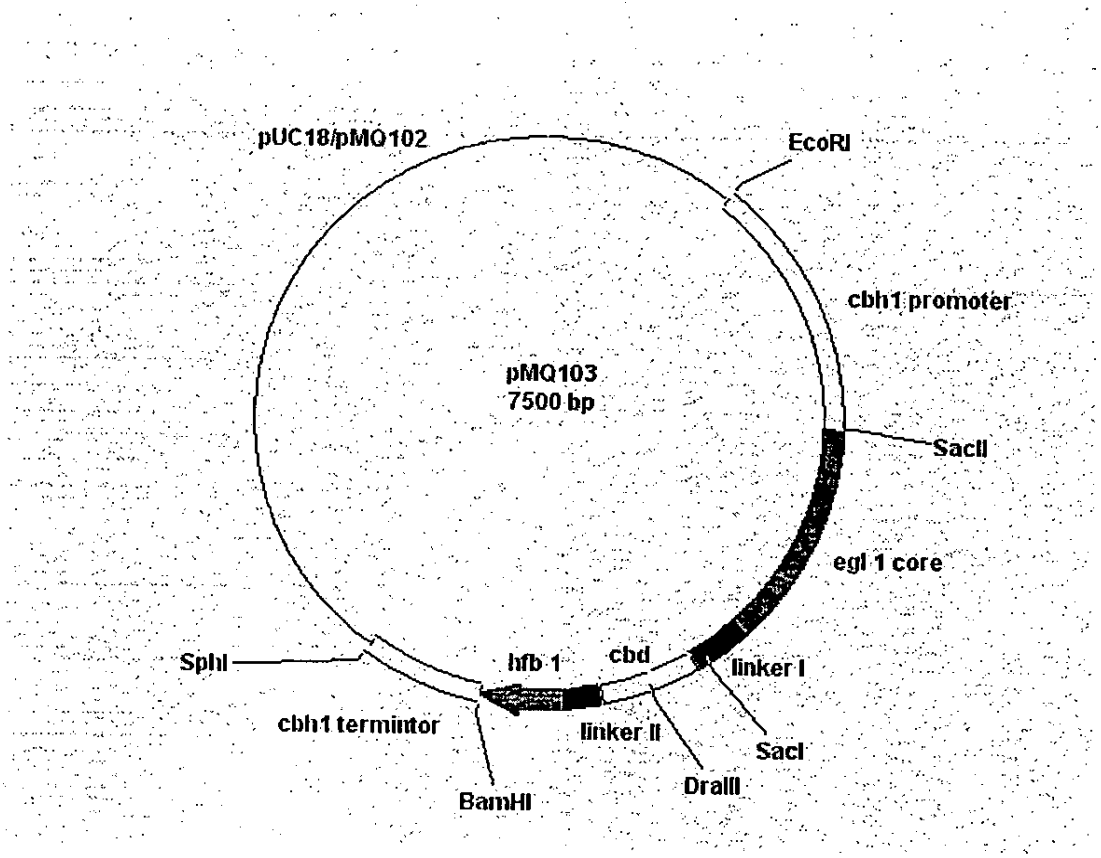

pUC18/pMQ102
EcoRI
cbh1 promoter
pMQ103
7500 bp
SacII
egl 1 core
SphI
hfb 1
cbd
linker I
cbh1 terminator
linker II
SacI
DraIII
BamHI

Fig 1.

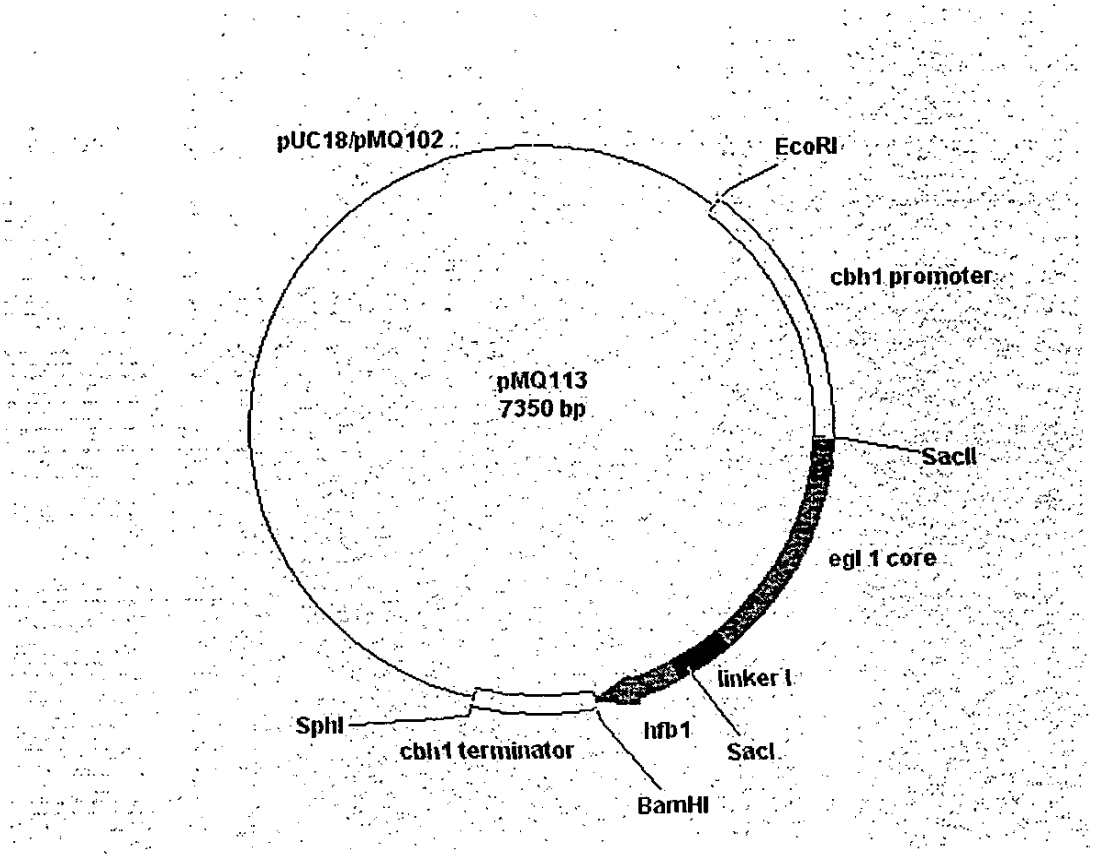

pUC18/pMQ102
EcoRI
cbh1 promoter
pMQ113
7350 bp
SacII
egl 1 core
linker I
SphI
hfb1
SacI
cbh1 terminator
BamHI

Fig 2.

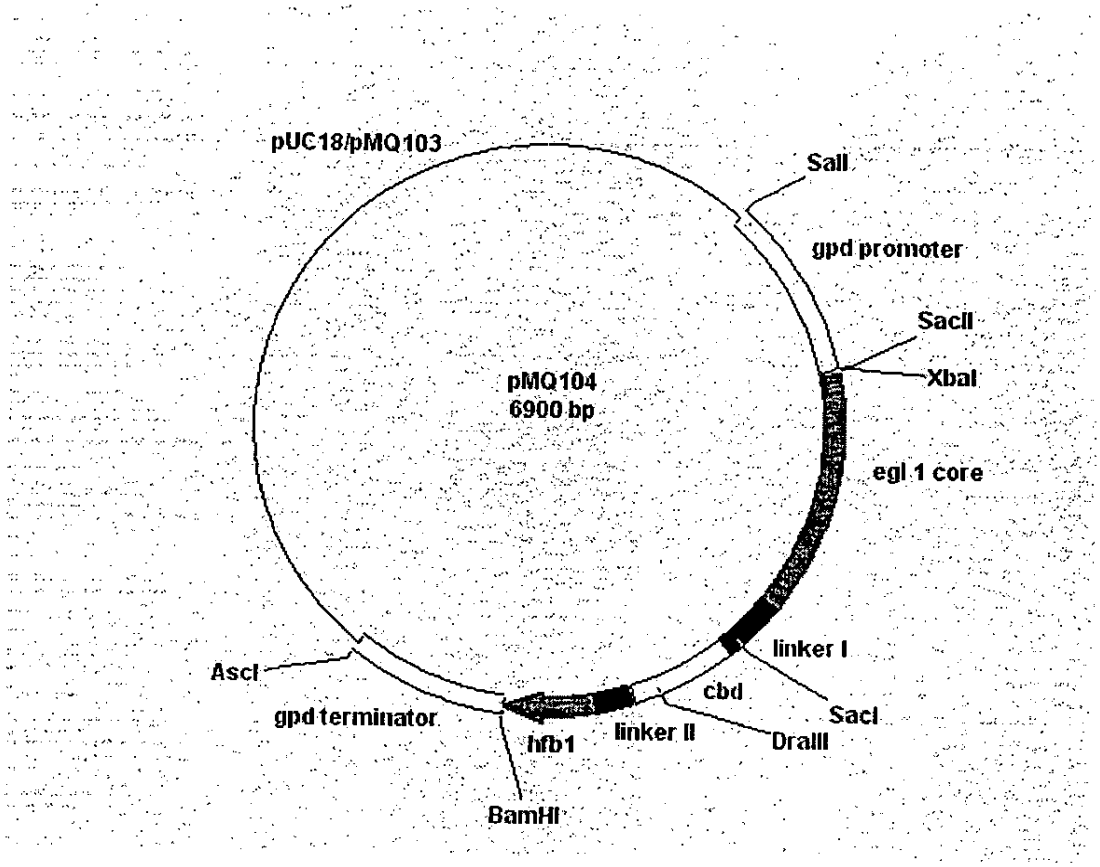

pUC18/pMQ103
SalI
gpd promoter
SacII
XbaI
pMQ104
6900 bp
egl 1 core
linker I
cbd
SacI
DraIII
linker II
hfb1
AscI
gpd terminator
BamHI

Fig 3.

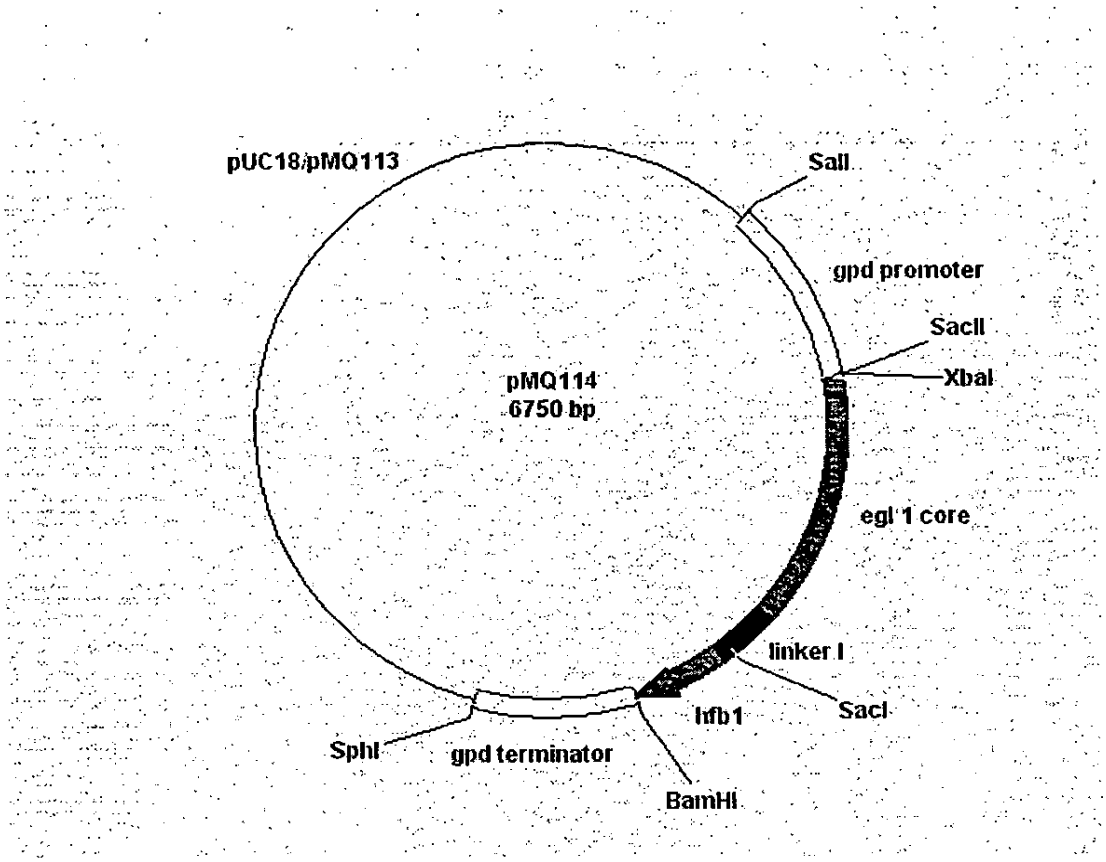

pUC18/pMQ113
SalI
gpd promoter
SacII
XbaI
pMQ114
6750 bp
egl 1 core
linker I
SacI
lifb1
SphI
gpd terminator
BamHI

Fig 4.

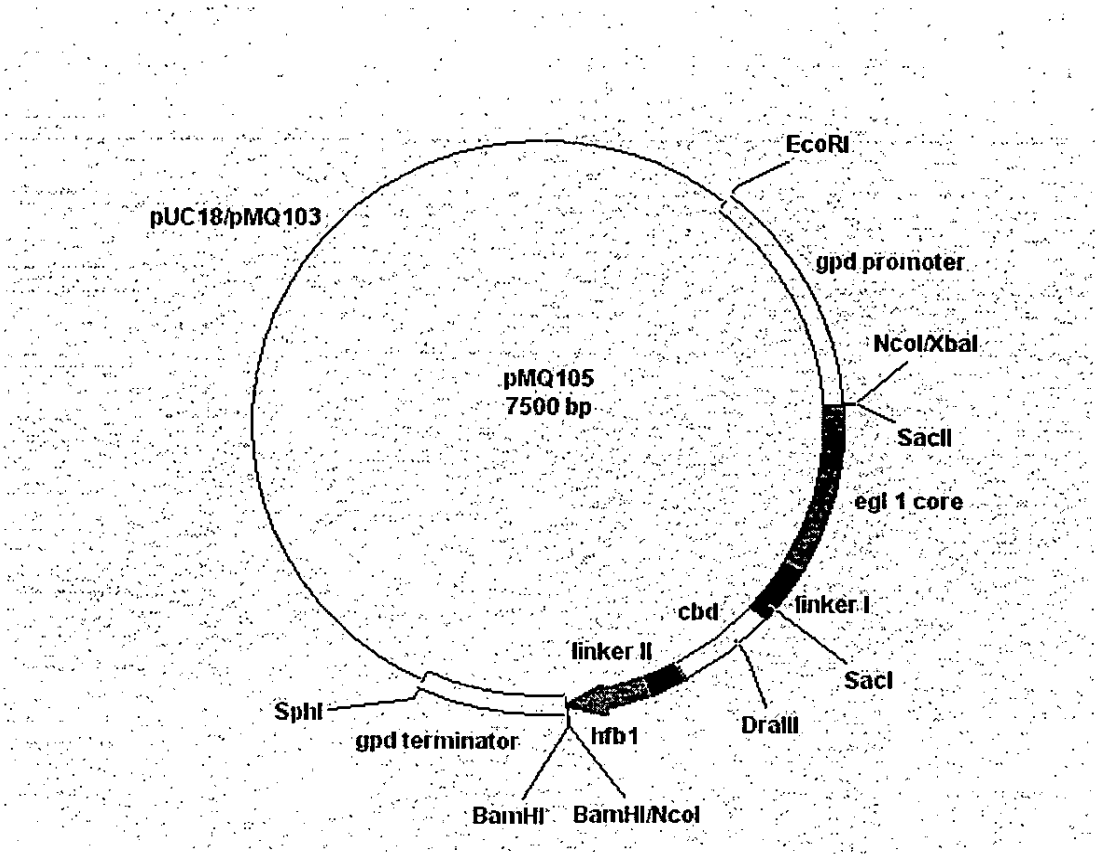

pUC18/pMQ103
pMQ105
7500 bp
EcoRI
gpd promoter
NcoI/XbaI
SacII
egl 1 core
linker I
SacI
DraIII
cbd
linker II
hfb1
BamHI
BamHI/NcoI
SphI
gpd terminator

Fig 5.

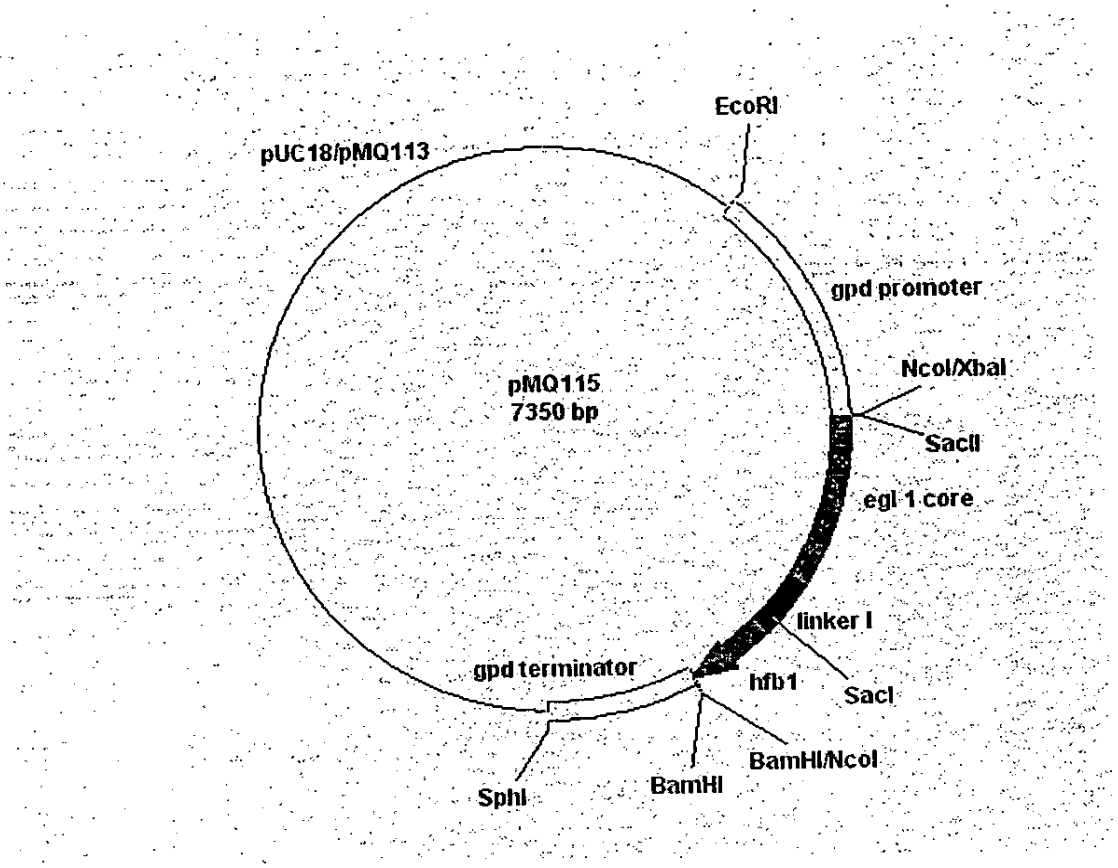

EcoRI
pUC18/pMQ113
gpd promoter
NcoI/XbaI
pMQ115
7350 bp
SacII
egl 1 core
linker I
gpd terminator
hfb1
SacI
BamHI/NcoI
SphI
BamHI

Fig 6.

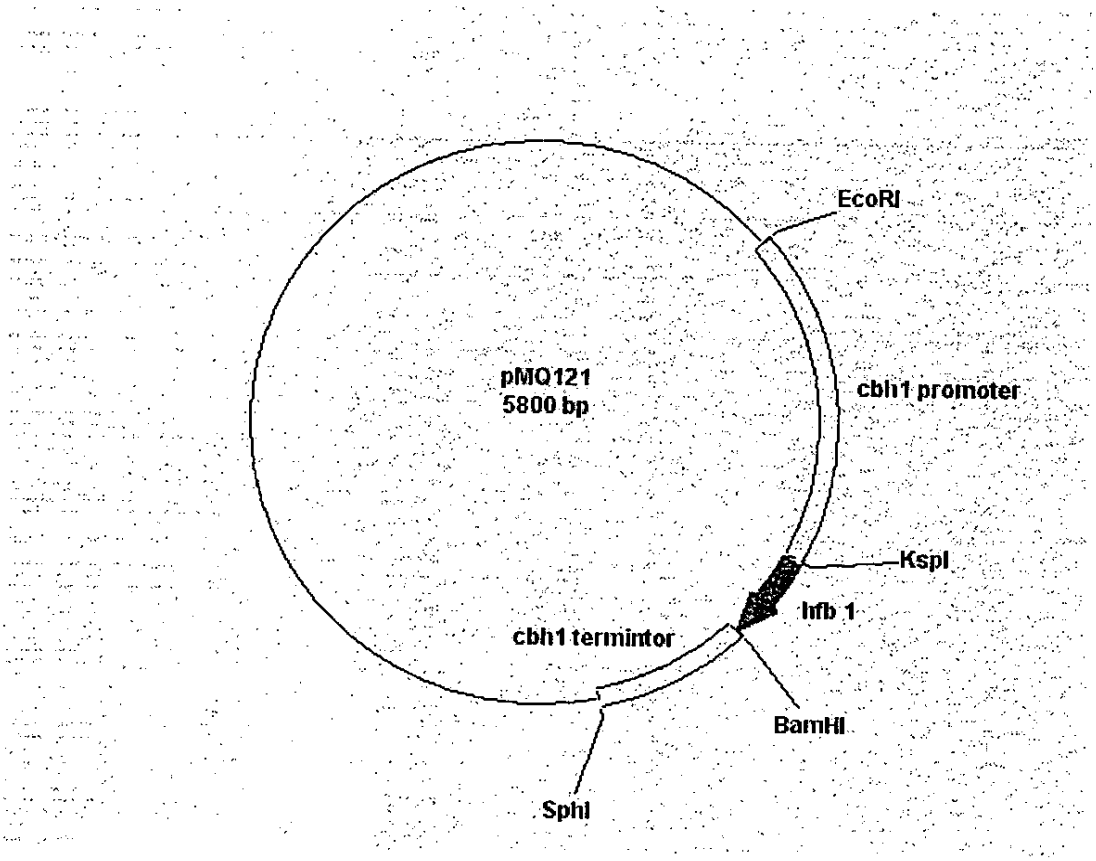

EcoRI
pMQ121
5800 bp
cbh1 promoter
KspI
hfb 1
cbh1 termintor
BamHI
SphI

Fig 7.

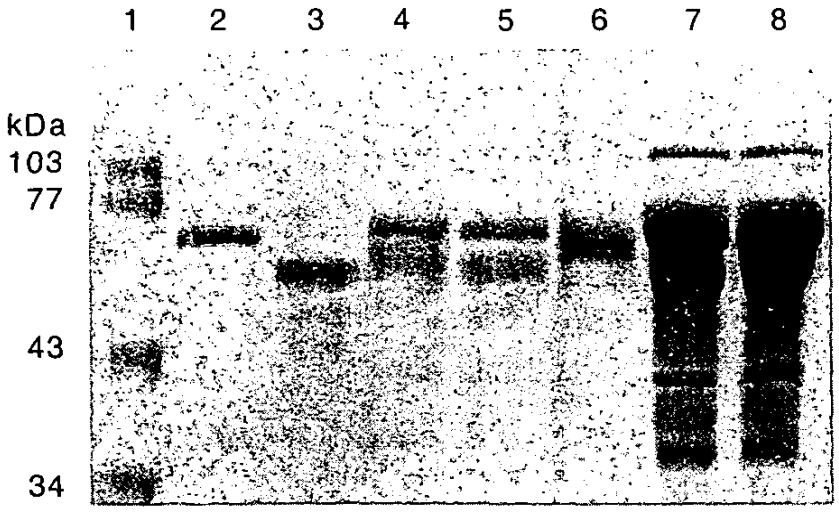


Figure 8. Coomassie stained 10 % SDS-PAGE of the partitioning of EGIcore-HFBI fusion protein in two-phase separation using 5 % of the detergent C12-C18EO5. Lane 1, Molecular weight marker; Lane 2, Purified CBHI (4 µg); Lane 3, Purified EGI (4 µg); Lane 4, 1/10 diluted VTT-D-98691 cellulose-based culture filtrate; Lanes 5 and 6, 1/10 diluted bottom phase and detergent phase (top phase), respectively, after separation of VTT-D-98691 culture filtrate with 5% detergent; Lane 7, Non-diluted bottom phase; Lane 8, Non-diluted VTT-D-98691 cellulose culture filtrate.

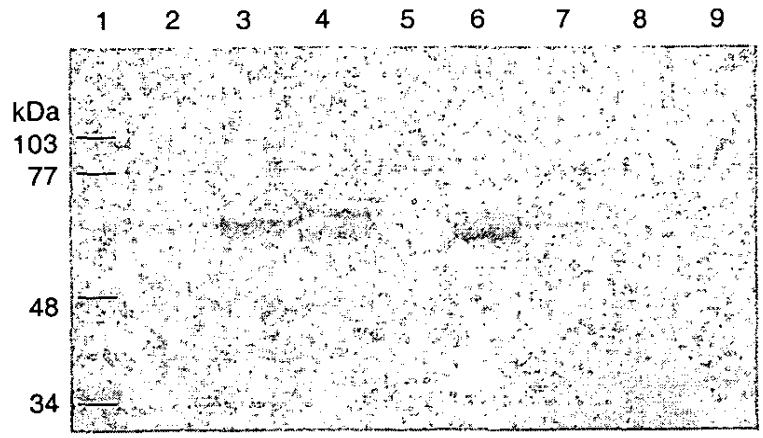


Figure 9. Western analysis of the partitioning of EGIcore-HFBI fusion protein in two-phase separation by using different concentrations of the detergent C12-C18EO5. Fusion proteins were detected with anti-HFBI antibodies. Lane 1, Molecular weight marker; Lane 2, Purified EGI; Lane 3, VTT-D-98691 cellulose culture filtrate; Lanes 4 and 5, Detergent phase (top phase) and bottom phase, respectively, after separation of VTT-D-98691 culture filtrate with 5% detergent; Lane 6, Same as lane 3, except 2 % detergent was used; Lane 7, Same as lane 4, except 2 % detergent was used; Lane 8, Purified EGI; Lane 9:, Purified CBHI.

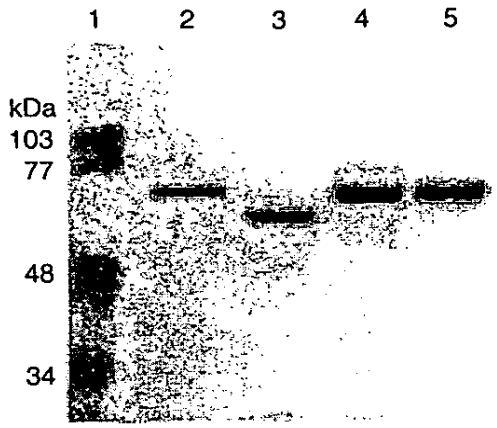


Figure 10. Coomassie stained 10 % SDS-PAGE showing further purification of EGIcore-HFBI fusion protein from the endogenous CBHI when the top phase was re-extracted with 2 % detergent. Lane 1, Molecular weight marker; Lane 2, Purified CBHI (4 μg); Lane 3, Purified EGI (4 μg); Lane 4, Detergent phase (top phase) after first extraction; Lane 5, Detergent phase (top phase) after second extraction.

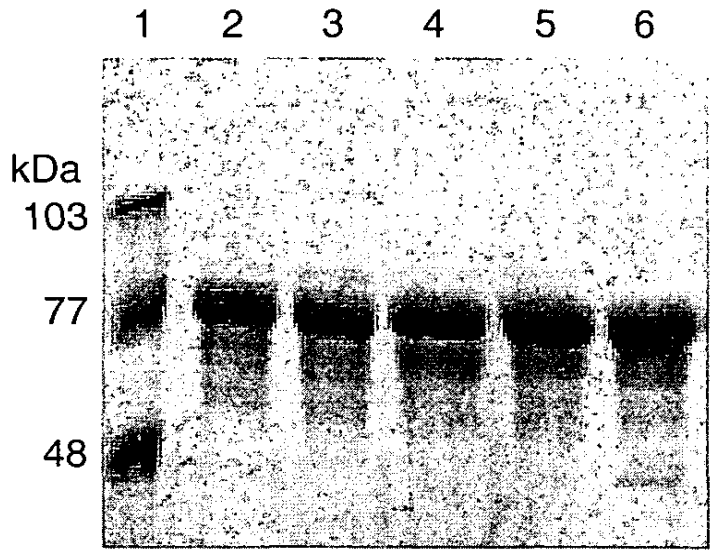


Figure 11. Coomassie stained 10 % SDS-PAGE analysis of the EGI-HFBI protein when treated with thrombin. Lane 1, Molecular weight marker; Lane 2, EGI-HFBI (1 mg/ml) treated 72 h with 3 U of thrombin at 24°C; Lane 3, Same as lane 2, except no thrombin was added; Lane 4, EGI-HFBI (1 mg/ml) treated 48 h with 9 U of thrombin at 36°C; Lane 5, Same as lane 4, except no thrombin was added; Lane 6, Same as lane 5, except no incubation at 36°C.

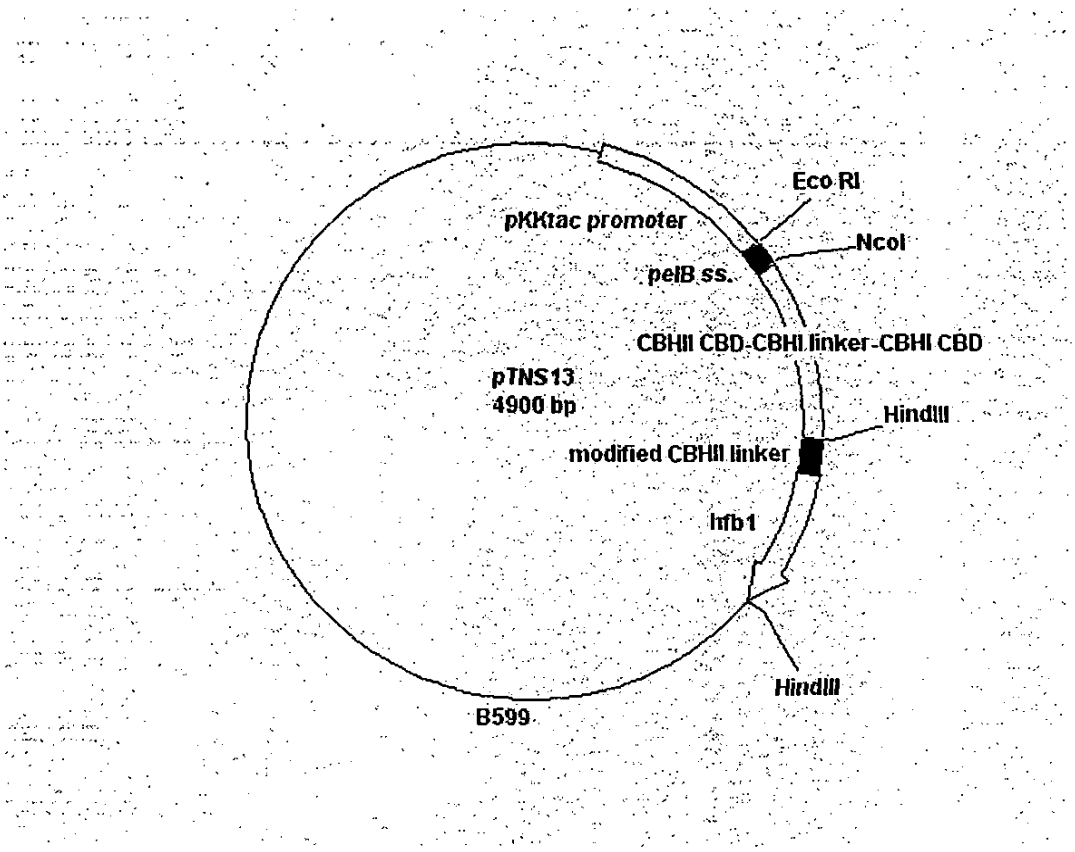
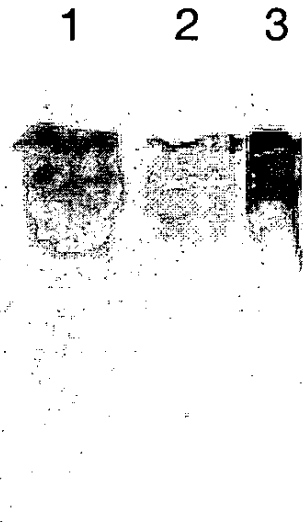

Eco RI
pKKtac promoter
NcoI
pelB ss.
CBHII CBD-CBHI linker-CBHI CBD
pTNS13
4900 bp
HindIII
modified CBHII linker
hfb1
HindIII
B599

Fig 12.

Figure 13. Western analysis of the partitioning of dCBD-HFB1 fusion protein in two-phase separation using 5 % of the detergent C12-C18EO5. Fusion protein was detected with anti-HFBI antibody. Lane 1, Four times concentrated culture filtrate; Lane 2, Four times concentrated bottom phase; Lane 3, Top phase.



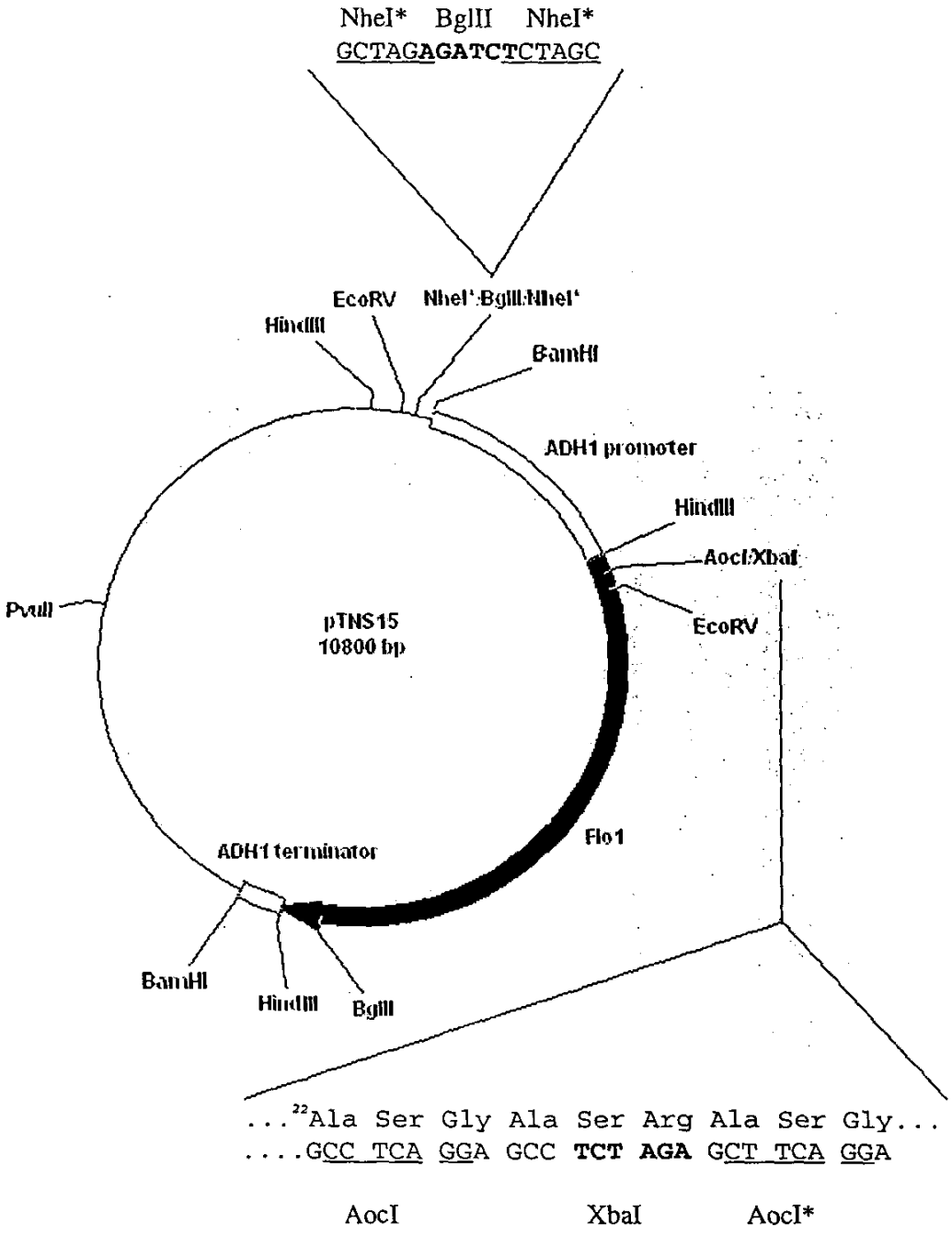


Fig 14.

A non-functional restriction site is indicated with an asterisk.

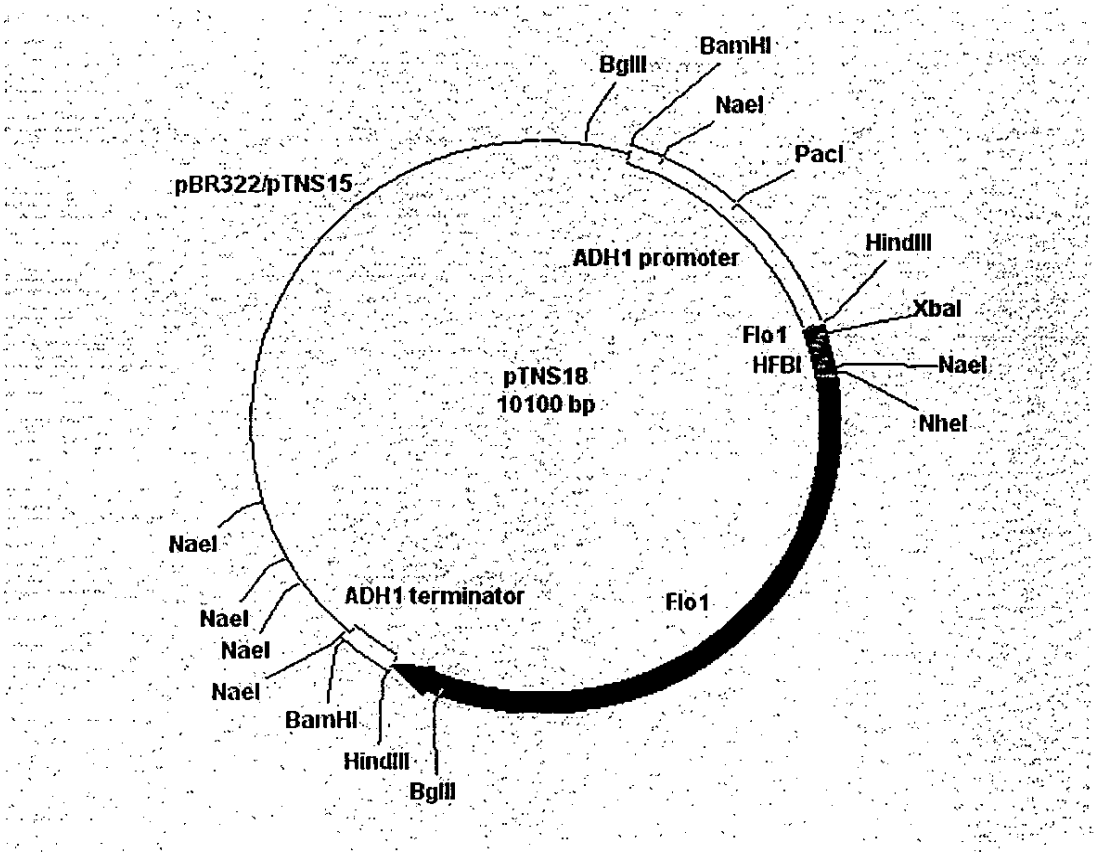

BglII
BamHI
NaeI
PacI
pBR322/pTNS15
ADH1 promoter
HindIII
XbaI
Flo1
HFBI
NaeI
NheI
pTNS18
10100 bp
NaeI
NaeI
NaeI
NaeI
ADH1 terminator
Flo1
BamHI
HindIII
BglII

Fig 15.

Figure 16